Synchronization of stochastic mean field networks of Hodgkin-Huxley neurons with noisy channels.

Mireille Bossy*, Joaquín Fontbona† and Héctor Olivero‡

December 27, 2018

Abstract

In this work we are interested in a mathematical model of the collective behavior of a fully connected network of finitely many neurons, when their number and when time go to infinity. We assume that every neuron follows a stochastic version of the Hodgkin-Huxley model, and that pairs of neurons interact through both electrical and chemical synapses, the global connectivity being of mean field type. When the leak conductance is strictly positive, we prove that if the initial voltages are uniformly bounded and the electrical interaction between neurons is strong enough, then, uniformly in the number of neurons, the whole system synchronizes exponentially fast as time goes to infinity, up to some error controlled by (and vanishing with) the channels noise level. Moreover, we prove that if the random initial condition is exchangeable, on every bounded time interval the propagation of chaos property for this system holds (regardless of the interaction intensities). Combining these results, we deduce that the nonlinear McKean-Vlasov equation describing an infinite network of such neurons concentrates, as time goes to infinity, around the dynamics of a single Hodgkin-Huxley neuron with chemical neurotransmitter channels. Our results are illustrated and complemented with numerical simulations.

Key words: Hodgkin-Huxley neurons, synchronization of neuron networks, mean-field limits, propagation of chaos, stochastic differential equations.

AMS Subject classification: 60H99, 60K35, 82C22, 82C32, 92B20, 92B25.

1 Introduction

The dynamics of a neuron's voltage is the result of the passage of ions through its membrane. This ion flux takes place through specific proteins which act as gated channels. According to the Hodgkin-Huxley model of a nerve neuron [29], the coupled behavior of the voltage of the neuron V_t, with the proportions m_t, h_t and n_t of open channels of the different ions involved in this process (respectively activation Sodium channels, deactivation Sodium channels and activation Potassium channels), can be described by the following system of ordinary differential equations:

$$\begin{aligned} V_t &= V_0 + \int_0^t F(V_s, m_s, n_s, h_s) ds \\ x_t &= x_0 + \int_0^t \rho_x(V_s)(1 - x_s) - \zeta_x(V_s) x_s \, ds \end{aligned} \tag{1.1}$$

where, here and in the sequel, x generically represents the m, n, h components and $F : \mathbb{R} \times [0,1]^4 \to \mathbb{R}$, defined by

$$F(V, m, n, h) = I - g_{\mathrm{K}} n^4 (V - V_{\mathrm{K}}) - g_{\mathrm{Na}} m^3 h (V - V_{\mathrm{Na}}) - g_{\mathrm{L}}(V - V_{\mathrm{L}}), \tag{1.2}$$

*INRIA Sophia Antipolis Méditerranée, France, Mireille.Bossy@inria.fr

†Center for Mathematical Modeling, University of Chile. fontbona@dim.uchile.cl. Supported by CMM-Basal Conicyt AFB 170001, Nucleo Milenio NC 130062 and Fondecyt Grant 1150570

‡CIMFAV, Facultad de Ingeniería, Universidad de Valparaíso, Chile. hector.olivero@uv.cl. Partially supported by Nucleo Milenio NC 130062, Beca Chile Postdoctorado and Fondecyt Postdoctoral Grant 3180777.

This research was partially supported by the supercomputing infrastructure of NLHPC (ECM-02), Conicyt.

represents the effect on the voltage of the ionic channels and of an external current I (assumed constant for simplicity). The rate functions ρ_x and ζ_x, originally considered in [29], have some generic form given in (2.3) and (2.4) below; see also Figure 2 and Tables 1 and 2 for their shape and for biologically meaningful values of the parameters. We refer the reader to Ermentrout and Terman [17] for a concise discussion on the Hodgkin-Huxley (HH in the sequel) model and its deduction, as well as for general background on mathematical models.

From a mathematical point of view, system (1.1) defines a rich dynamical system, the properties of which has been extensively studied. As an example of its various possible behaviors, Figure 3 below illustrates different possible responses of system (1.1) to the value of the input current I (no oscillations, oscillations of various types, damping) all other parameters of the model being fixed. See e.g. [17] and Izhikevich [31], and references therein for detailed accounts on dynamical properties of (1.1) and related neuron models. Lower dimensional dynamics have also been proposed as simpler alternatives to (1.1), the most important ones being the FitzHugh-Nagumo model (FitzHugh [19], Nagumo et al. [42]) and the Morris-Lecar model (Morris and Lecar [41]). These are able to reproduce some of the dynamical features of the HH system (1.1) and are easier to study from the mathematical point of view, but they are less realistic regarding some of its relevant features.

A different approach to model the electric activity of neurons are integrate-and-fire models, introduced in Lapicque [36]. In these models the electric potential evolves according to some ordinary differential equation until it reaches a certain fixed threshold; the neuron then emits a potential spike and the voltage is reset to some reference value, from which its evolution restarts following the same dynamics. We refer the reader to Burkitt [12], [13] for a review of this class of models.

In the last decade, there has been an increasing interest of the mathematical and computational neuroscience communities in understanding the role of stochasticity in neurons' dynamics, as well as in mathematical models for it. We refer the reader to Goldwyn et al. [25] and to Goldwyn and Shea-Brown [26] for a discussion on different ways in which randomness might be introduced in the HH model, their biological interpretation and their pertinence. See also [17, Chapter 10] for general background on this issue. Classically, random models arise in the form of finite Markov chains describing a discrete number of open gates which approximate the ion channel dynamics, or by directly introducing Gaussian additive or multiplicative white noise (that is, a Brownian motion or a stochastic integral with respect to it) in the voltage or ion channels dynamics in (1.1). More recently, hybrid (also called piecewise deterministic) Markov processes have also been proposed as microscopic counterparts of the HH or other deterministic models. In this setting, the channel variables are replaced by discrete continuous time processes whose jump rates depend on the voltage, while keeping a continuous description for the latter, see Austin [2], Pakdaman et al. [44] and references therein. We also refer the reader to Bossy et al. [10], Dangerfield et al. [15], Wainrib [51] and Sacerdote and Giraudo [48] for further discussion on stochastic models in this context, the latter one in the case of integrate-and-fire models.

In the present work we consider stochastic versions of the HH model (1.1) which arise as diffusive scaling limits of hybrid models of the type studied in [2] and [44]. More precisely, we are interested in networks of N such neurons in mean field interaction, which can be described by a system of stochastic differential equations of the form:

$$
\begin{aligned}
V_t^{(i)} &= V_0^{(i)} + \int_0^t F(V_s^{(i)}, m_s^{(i)}, n_s^{(i)}, h_s^{(i)})ds \\
&\qquad - \int_0^t \frac{1}{N}\sum_{j=1}^N J_{\mathrm{E}}(V_s^{(i)} - V_s^{(j)}) - \frac{1}{N}\sum_{j=1}^N J_{\mathrm{Ch}} y_s^{(j)}(V_s^{(i)} - V_{\mathrm{rev}})ds, \\
x_t^{(i)} &= x_0^{(i)} + \int_0^t b_x(V_s^{(i)}, x_s^{(i)})ds + \int_0^t \sigma_x(V_s^{(i)}, x_s^{(i)})dW_s^{x,i}, \quad x = m, n, h, y.
\end{aligned}
$$

Here, $(W^{x,i} : i = 1, \dots, N, x = m, n, h, y)$ are one dimensional Brownian motions and the interaction between neurons account for the effect of electrical and chemical synapses (the biological interpretation of the interactions terms and in particular of the variables $y^{(i)}$ is given in next section). We refer to equation (2.1) below for the explicit form of the system we will consider, and for Hypothesis 2.1 for our assumptions on its coefficients.

The collective behavior of neurons, and the way it emerges from their individual features and synaptic activity, is indeed a central question in neuroscience. In particular, considerable efforts have been devoted to understanding synchronization of neurons, an ubiquitous phenomenon seemingly related to the generation of rhythms (such as the respiratory one or the heartbeat) but also to more complex neurologic functionalities. For instance, at the brain level, synchronization has been connected to memory formation, see Axmacher et al. [3], but also to disorders such as epileptic seizures, see Jiruska et al. [32]. Since the neuroscience literature on this topic is huge, it is not our intention to thoroughly comment on it here, and we refer the reader to [17, Chapters 8,9], [31, Chapter 10] for further discussion on biological roles of neuron synchronization, and mathematical approaches to it. For a broad perspective on synchronization, we also refer the reader to Pikovsky et al. [47].

In this paper we are interested in synchronization due to a strong enough coupling between neurons. This phenomenon differs from synchronization owed to common noise addressed e.g. in Marella and Ermentrout [38] (where uncoupled oscillators subject to a common noise are observed to get synchronized) or in Pikovsky [46] (where synchronization results from the action of a random forcing term). In our case, noise is unshared by the interacting neurons and does therefore not contribute to their synchronization; indeed, it actually prevents the perfect asymptotic synchronization of the network. (This phenomenon might be compared to noise-induced deviations from stable cycles in noisy oscillators, see [5, 6] for a large deviations approach to that problem.) Moreover, we will understand and quantify synchronization in terms of the empirical variance of the system of neurons and of its dissipation, an approach which does not rely on the stability properties of individual neuron dynamics nor, in particular, on the existence of some oscillatory limiting behavior.

In the case of interacting oscillators, a central mathematical tool is phase reduction. Introduced by Kuramoto [35], it is based on the idea that stable periodic solutions of a nonlinear oscillator can be parametrized by its phase in the limit cycle. Kuramoto's model has proved useful to understand synchronization mechanisms of simple coupled oscillators (see [17] and [31]), or for ensembles of population of neurons with intrinsic and extrinsic noise (see Bressloff and Ming Lai [11] and the references therein), or even in the limiting case of infinite oscillators with noise and mean field interaction (see Bertini et al. [7] and the discussion in [51, Chapter 4]). However, to our knowledge, applications of these ideas to networks of HH-type neurons have so far been restricted to small deterministic networks and "weak coupling" regimes (see e.g. Hansel and Mato [27] and Hansel et al. [28]). We refer the reader to Ostojic et al. [43] for synchronization results in the case of integrate-and-fire networks.

A related question is the asymptotic behavior of networks when the number of neurons tends to infinity. In that sense, networks of N neurons in mean field interaction, in which every neuron experiences a pairwise interaction of strength-order $1/N$ with each other, provide a mathematically tractable (though not completely realistic) framework to address this question. Indeed, in a mean field network, the evolutions of finitely many neurons are expected to become independent as N goes to infinity, a property known as propagation of chaos. In the case of exchangeable particles, this is equivalent to the convergence of the dynamics of the empirical law of the system to some deterministic flow of probability laws, typically described by a nonlinear McKean-Vlasov partial differential equation (also termed "mean field" equation in this context); see Méléard [39], Sznitman [49] for background on propagation of chaos. We refer the reader to e.g. Faugeras et al. [18] for formal derivations of mean field equations for multi type population networks of integrate-and-fire neurons, and respectively to Delarue et al. [16] and Perthame and Salort [45] for probabilistic and PDE approaches to the global solvability of that equation (which can in principle have explosive solutions) when the interaction is small. See also Fournier and Löcherbach [21] for further recent results on propagation of chaos for integrate and fire models. The propagation of chaos for mean field networks of neurons described by stochastic differential equations, including stochastic, multi type HH and FitzHugh-Nagumo networks, has been addressed in Baladron et al. [4], and then rigorously established in Bossy et al. [10]. We also refer the reader to Mischler et al. [40] for the mean field description of a network of FitzHugh-Nagumo neurons.

In this present work, we establish that, under strong enough electrical connectivity of the network (i.e. large enough J_{E}), the N neurons get synchronized, up to an error proportional to the channels' noise level σ^2, at an exponential rate which is independent of N. Moreover, we exhibit a deterministic single-neuron dynamics which is "mimicked", as time goes to infinity, by every neuron of the system (2.1), over short enough moving time-windows, up to an error that vanishes with σ^2 and N^{-1}. As far as we know, this is a first mathematical result which establishes the synchronization of large networks of neurons. We also establish the propagation of chaos for system (2.1), or its convergence to solutions to a McKean-Valsov equation, for arbitrary parameters of the model (and for slightly more general coefficients than in [10] in the single population case). This allows us to transfer our synchronization results to the limiting PDE, which can be understood as the description of an infinite network of neurons. Our theoretical results will be complemented with simulations, which in particular point out that synchronization phenomena might also hold for small electrical connectivity and even for pure chemical connectivity ($J_{\mathrm{E}} = 0$).

The remainder of the paper is organized as follows. In Section 2, we detail the model we consider and state precisely our main results. Section 3 is devoted to numerical experiments, both to illustrate our theoretical statements and to explore related phenomena in mean field settings not covered by them (like multi-type neuron populations or chemical-only synapses). In Section 4 we a discuss possible improvements and extensions of our results, and some open questions. The mathematical proofs of our results are given in the Appendix sections.

2 Model and main results

We start by briefly recalling how chemical and electrical synapses in networks of neurons are modeled (we follow [17, Chapter 7] which we also refer to for further background on synaptic channels).

In chemical synapses, a neurotransmitter is released to the intercellular media (technically the synaptic cleft), from a pre-synaptic neuron to the post-synaptic one through synaptic channels, which are voltage-gated just as ion channels are. With each pre-synaptic neuron we can thus associate a new variable y in $[0, 1]$ which represents its proportion of open synaptic channels at each time. The dynamics of this variable can be modeled in a similar way as those of ion channels, that is, in terms of certain rate functions ρ_y and ζ_y depending on the membrane potential V of that same neuron, and on some parameters (see (2.5)). The choice of these parameters determines the characteristic (inhibitory or excitatory) of the chemical synapse. Hence, in a fully connected network of N similar neurons, chemical synapses coming from a pre-synaptic neuron j should induce on the voltage $V^{(i)}$ of the post-synaptic neuron i an instantaneous variation at time t of

$$-\frac{J_{\mathrm{Ch}}}{N} y_t^{(j)} (V_t^{(i)} - V_{\mathrm{rev}}),$$

where $y_t^{(j)}$ is the proportion of open synaptic channels of neuron j, $J_{\mathrm{Ch}} \geq 0$ is a constant representing the chemical conductance of the network and V_{rev} is a reference potential. The factor $\frac{1}{N}$ is introduced in order that the contribution of each incoming synapse to the neuron i has similar weight, which corresponds to a global interaction of mean field type.

On the other hand, the interior of one neuron can be directly connected with another neuron's one through an intercellular channel called gap junction, which allows the constant flow of ions between them, as a result of their possibly different potentials. We thus may assume that pre-synaptic neuron j contributes to the variation of the voltage of post-synaptic neuron i by the amount

$$-\frac{J_{\mathrm{E}}}{N} (V_s^{(i)} - V_s^{(j)}),$$

where $J_{\mathrm{E}} \geq 0$ is the electrical conductance (that can be thought of as a measure of the connectivity of the network) and the factor $\frac{1}{N}$ appears by similar reasons as before. Connections of this type are termed electrical synapses and are less frequent than chemical ones; on the other hand, they transmit information faster. (See also Hormuzdi et al. [30] for a deeper discussion on electrical synapses.)

In all the sequel, for each fixed N we consider a stochastic process $X = (X^{(1)}, \ldots, X^{(N)})$ valued in $(\mathbb{R}^5)^N$, with coordinates $X_t^{(i)} = (V_t^{(i)}, m_t^{(i)}, n_t^{(i)}, h_t^{(i)}, y_t^{(i)})$ given for $i = 1, \ldots, N$ and $t \geq 0$ by the solution of the system of stochastic differential equations:

$$\begin{aligned} V_t^{(i)} &= V_0^{(i)} + \int_0^t F(V_s^{(i)}, m_s^{(i)}, n_s^{(i)}, h_s^{(i)}) ds \\ &\quad - \int_0^t \frac{1}{N} \sum_{j=1}^N J_{\mathrm{E}} (V_s^{(i)} - V_s^{(j)}) - \frac{1}{N} \sum_{j=1}^N J_{\mathrm{Ch}} y_s^{(j)} (V_s^{(i)} - V_{\mathrm{rev}}) ds, \\ x_t^{(i)} &= x_0^{(i)} + \int_0^t \rho_x(V_s^{(i)})(1 - x_s^{(i)}) - \zeta_x(V_s^{(i)}) x_s^{(i)} ds + \int_0^t \sigma_x(V_s^{(i)}, x_s^{(i)}) dW_s^{x,i}, \end{aligned} \tag{2.1}$$

where $(W^{x,i} : i \in \mathbb{N}, x = m, n, h, y)$ are independent one dimensional Brownian motions independent of X_0 and F is defined in (1.2). Notice that, for notational simplicity, the dependence of system (2.1) on N is omitted. Throughout this work, we will additionally make the following assumptions on system (2.1):

Hypothesis 2.1. *1. For $x = m, n, h$ and y, ρ_x and ζ_x are strictly positive, locally Lipschitz continuous functions defined on $\mathbb{R}$.*

2. For $x = m, n, h$ and y, functions $\sigma_x : \mathbb{R}^2 \to \mathbb{R}$ are given by

$$\sigma_x(v, z) = \sigma\sqrt{|\rho_x(v)(1-z) + \zeta_x(v)z|}\chi(z), \tag{2.2}$$

with $\chi : \mathbb{R} \to [0, 1]$ a Lipschitz continuous function with support contained in $[0, 1]$ and $\sigma \geq 0$.

3. One has $(m_0^{(i)}, n_0^{(i)}, h_0^{(i)}, y_0^{(i)}) \in [0, 1]^4$ a.s.

These assumptions cover, for parameters $a_r^x, a_d^x > 0$, functions of the form

$$\rho_x(V) = \frac{a_r^x(V - V_r^x)}{1 - \exp\left(-\lambda_r^x(V - V_r^x)\right)}, \quad \zeta_x(V) = a_d^x \exp\left(-\lambda_d^x(V - V_d^x)\right), \tag{2.3}$$

for $x = m, n$, and

$$\rho_h(V) = a_r^h \exp\left(-\lambda_r^h(V - V_r^h)\right), \quad \zeta_h(V) = \frac{a_d^h}{1 + \exp\left(-\lambda_d^h(V - V_d^h)\right)}, \tag{2.4}$$

considered in the original HH model [29], as well as functions

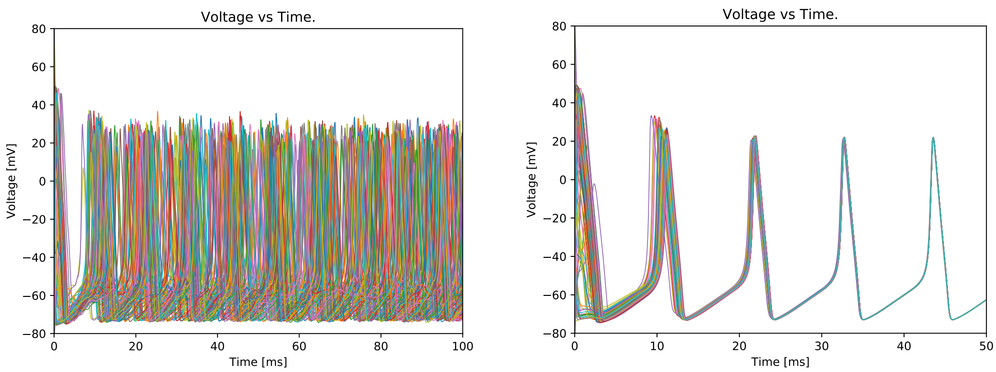
$$\rho_y(V) = \frac{a_r^y T_{\max}}{1 + \exp\left(-\lambda(V - V_T)\right)}, \quad \zeta_y(V) = a_d^y \tag{2.5}$$

associated with synaptic channels in [17, Chapter 7]. Diffusion coefficients σ_x defined in terms of the functions ρ_x and ζ_x as in (2.2) have been considered in [10], and arise naturally in diffusive scaling limits of the hybrid models studied in [44].

Observe that for functions ρ_h, ζ_m and ζ_n as above, the coefficients of system (2.1) do not satisfy classic conditions for wellposedness. This well-posedness will be proved in Lemma A.1 below, relying on results in [10] that ensure that under Hypothesis 2.1 the processes $(m_t^{(i)}, n_t^{(i)}, h_t^{(i)}, y_t^{(i)})$ remain in $[0, 1]^4$. Notice that the absolute value in (2.2) can then be removed.

Synchronization

By synchronization we will understand the dissipation of the empirical variance of the network (2.1) as time goes by. In Figure 1 we show two extreme situations in this regard. On the left, for a small interaction parameter J_{E} and noise $\sigma \neq 0$ we observe a chaotic behavior resulting in an empirical variance of constant order in time. On the right, for large J_{E} and $\sigma = 0$ we observe the fast emergence of a coherent evolution implying the dissipation of the empirical variance. Our main result, Theorem 2.3 will provide a quantitative picture of this behavior with respect to noise level σ and the size of the network N, for large enough connectivity J_{E}. Our results require the following additional assumption:



(a) Small interaction parameter and noise. (b) Large interaction parameter and no noise.

Figure 1: Two typical situations of the evolution of the network (2.1).

Hypothesis 2.2. *Hypothesis 2.1 holds and moreover:*

4. *The parameter g_L in (1.2) (termed "leak conductance") is strictly positive.*
5. *There exists a constant $V_0^{max} > 0$ not depending on N such that*

$$\sup_{i=1,\dots,N} |V_0^{(i)}| \leq V_0^{max} \quad a.s.$$

We also need to introduce notation for some empirical means, namely

$$\bar{V}_t^N = \frac{1}{N}\sum_{i=1}^N V_t^{(i)}, \quad \bar{X}_t^N = \frac{1}{N}\sum_{i=1}^N X_t^{(i)} \quad \text{and} \quad \bar{x}_t^N = \frac{1}{N}\sum_{i=1}^N x_t^{(i)} \text{ for } x = m, n, h, y.$$

Moreover, for each $N \geq 1$ and $t_1 \geq 0$ we let $(\widehat{X}_t^{N,t_1} : t \geq t_1) = \Big((\widehat{V}_t^N, \widehat{m}_t^N, \widehat{n}_t^N, \widehat{h}_t^N, \widehat{y}_t^N) : t \geq t_1\Big)$ denote the solution of the ordinary differential equation

$$\begin{aligned} \widehat{V}_t^N &= \widehat{V}_{t_1}^N + \int_{t_1}^t F(\widehat{V}_s^N, \widehat{m}_s^N, \widehat{n}_s^N, \widehat{h}_s^N) - J_{\text{Ch}}\widehat{y}_s^N(\widehat{V}_s^N - V_{\text{rev}})ds, \\ \widehat{x}_t^N &= \widehat{x}_{t_1}^N + \int_{t_1}^t \rho_x(\widehat{V}_s^N)(1-\widehat{x}_s^N) - \zeta_x(\widehat{V}_s^N)\widehat{x}_s^N ds, \quad x = m, n, h, y. \end{aligned} \tag{2.6}$$

with random initial condition

$$\widehat{X}_{t_1}^{N,t_1} = \bar{X}_{t_1}^N.$$

We are now is position to state our main result about synchronization of the system (2.1):

Theorem 2.3. *Assume Hypothesis 2.2 holds and that $(X_0^{(1)}, \dots, X_0^{(N)})$ is an exchangeable random vector.*

a) **Synchronization.** *There exist constants $J_E^0 > 0$, $C_{\zeta,\rho}^0 > 0$ and $\lambda^0 > 0$ not depending on $N \geq 1$, $\sigma \geq 0$ or X_0, and there exists a time $t_0 \geq 0$ not depending on $N \geq 1$ or $\sigma \geq 0$, such that for each $J_E > J_E^0$ the solution X of (2.1) satisfies, for every $t \geq t_0$ and each $i \in \{1, \dots, N\}$:*

$$\mathbb{E}\left(|X_t^{(i)} - \bar{X}_t^N|^2\right) \leq \mathbb{E}\left(|X_{t_0}^{(i)} - \bar{X}_{t_0}|^2\right)e^{-\lambda^0(t-t_0)} + \sigma^2\frac{C_{\zeta,\rho}^0}{\lambda^0}. \tag{2.7}$$

In particular, $\limsup_{t\to\infty} \mathbb{E}\left(|X_t^{(i)} - \bar{X}_t^N|^2\right) \leq \sigma^2 \frac{C_{\zeta,\rho}^0}{\lambda^0}$.

b) **Synchronized dynamics.** *Assume $J_E > J_E^0$. Then, there are constant $K_0, K_0' > 0$ depending only on the parameters of the voltage dynamics in (1.2) and, for each $\delta \geq 0$, constants $K_\delta, K_\delta' > 0$ depending on the coefficients in (2.1) and on δ (increasingly) but not on N, such that for every $t_1 \geq t_0$ and each $i \in \{1, \dots, N\}$:*

$$\begin{aligned} &\sup_{t_1 \leq t \leq t_1+\delta} \mathbb{E}\left(|X_t^{(i)} - \widehat{X}_t^{t_1,N}|^2\right) \\ &\quad \leq K_0 \wedge 2\left[\left(K_0' e^{-\lambda_0(t_1-t_0)} + \sigma^2\frac{C_{\zeta,\rho}^0}{\lambda^0}\right)(1+\delta K_\delta) + \delta K_\delta'\frac{\sigma^2}{N}C_{\zeta,\rho}^0\right]. \end{aligned} \tag{2.8}$$

Some remarks on this result are in order:

Remark 2.4.

(i) The constants $C_{\zeta,\rho}^0$ and t_0 depend explicitly on the coefficients of the system, with the latter possibly depending also on V_0^{max}. On the other hand, bounds for λ_0 and J_E^0 which do not depend on the initial data can also be obtained. The remaining constants are explicit and do not depend on the initial condition. No constant is claimed to be optimal.

(ii) The bound K_0 in Theorem 2.3 b) is deduced from global bounds (which we establish) on the voltage processes and their average, and its role is only to prevent the r.h.s. from growing arbitrarily with δ. The estimate becomes informative as $t_1 \to \infty$ for small enough $\sigma^2 > 0$, $\delta > 0$ and N^{-1}, and for any $\delta > 0$ and N if $\sigma^2 = 0$.

(iii) *Aside from the assumption that $g_L > 0$, Theorem 2.3 holds regardless of the values of the parameters of voltage dynamics F in (1.2), and in particular if the input current I is replaced by a uniformly bounded function or (suitably measurable) process.*

(iv) *The exchangeability assumption can be removed at the price of adding inside the expectations in (2.7) and (2.8) the averages over i.*

Mean field limit

We next address the question of the behavior of system (2.1) as $N \to \infty$. We need to introduce additional notation:

- We denote by $\mathcal{P}(\mathbb{R} \times [0,1]^4)$ the space of Borel probability measures on $\mathbb{R} \times [0,1]^4$ endowed with the weak topology, and by $\mathcal{P}_2(\mathbb{R} \times [0,1]^4)$ its subspace of probability measures with finite second moment, endowed with the Wasserstein distance $\mathcal{W}_2$. That is, for all $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R} \times [0,1]^4)$,

$$\mathcal{W}^2(\mu_1, \mu_2) = \inf_{\mu \in \Pi(\mu_1,\mu_2)} \int_{\mathbb{R}^5} |r_1 - r_2|^2 \mu(dr_1, dr_2),$$

  with $\Pi(\mu_1, \mu_2)$ the set of probability measures on $(\mathbb{R} \times [0,1]^4)^2$ with first and second marginals equal to μ_1 and μ_2 respectively. It is well known that the infimum is attained and that $\mathcal{W}_2$ defines a complete metric on $\mathcal{P}_2(\mathbb{R} \times [0,1]^4)$ inducing the weak topology, strengthened with the convergence of second moments (see Villani [50] for the relevant properties of Wasserstein metrics).
- Elements of $\mathbb{R} \times [0,1]^4$ describing the state space of a single neuron's dynamics will be written in the form $(v, u) = (v, (u_m, u_n, u_h, u_y))$, with $u = (u_m, u_n, u_h, u_y) \in [0,1]^4$.
- We introduce the function $\Phi : (\mathbb{R} \times [0,1]) \times (\mathbb{R} \times [0,1]^4) \to \mathbb{R}$ given by

$$\Phi(w, z, v, u) = F(v, u_m, u_n, u_h) - J_{\text{E}}(v - w) - J_{\text{Ch}} z (v - V_{\text{rev}})$$

  and, for each channel type $x = m, n, h, y$, we define functions $b_x, a_x : \mathbb{R} \times [0,1]^4 \to \mathbb{R}$ by

$$\begin{aligned} b_x(v,u) &:= \rho_x(v)(1 - u_x) - \zeta_x(v) u_x \quad \text{and} \\ a_x(v,u) &:= (\rho_x(v)(1 - u_x) + \zeta_x(v) u_x) \chi(u_x) \end{aligned}$$

  (that is, $\sigma_x^2(v, u_x) = \sigma^2 a_x(v, u)$).
- Given $\mu \in \mathcal{P}_2(\mathbb{R} \times [0,1]^4)$ we write

$$\begin{aligned} \langle \mu^V \rangle &= \int_{\mathbb{R}^5} v \mu(dv, du_m, du_n, du_h, du_y) \in \mathbb{R}, \\ \langle \mu^x \rangle &= \int_{\mathbb{R}^5} u_x \mu(dv, , du_m, du_n, du_h, du_y) \in [0,1] \text{ for } x = m, n, h, y \text{ and} \\ \langle \mu \rangle &= (\langle \mu^V \rangle, (\langle \mu^x \rangle)_{x=m,n,h,y}) \in \mathbb{R} \times [0,1]^4. \end{aligned}$$

- Finally, with δ_x denoting the Dirac mass at $x \in \mathbb{R} \times [0,1]^4$, we write

$$\mu_t^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^{(i)}} \in \mathcal{P}_2(\mathbb{R} \times [0,1]^4) \tag{2.9}$$

  for the empirical measure of system (2.1) at time $t \geq 0$.

Theorem 2.5. *Assume Hypothesis 2.2 and moreover that for all $N \geq 1$, $(X_0^{(1)}, \dots, X_0^{(N)})$ are i.i.d. random vectors with (compactly supported) common law $\mu_0 \in \mathcal{P}(\mathbb{R} \times [0,1]^4)$ not depending on N.*

a) *For each $T > 0$, the process $(\mu_t^N : t \in [0,T])$ converges in law on $C([0,T]; \mathcal{P}_2(\mathbb{R} \times [0,1]^4))$, when N tends to ∞, to a deterministic and uniquely determined flow of probability measures $(\mu_t : t \in [0,T])$ having uniformly bounded compact support. Moreover $(\mu_t : t \geq 0)$ in $C(\mathbb{R}^+; \mathcal{P}_2(\mathbb{R} \times [0,1]^4))$ is a global solution (in the sense of distribution) of the non linear McKean-Vlasov Fokker Planck equation*

$$\partial_t \mu_t = \partial_v \left(\Phi(\langle \mu_t^V \rangle, \langle \mu_t^y \rangle, \cdot, \cdot) \mu_t \right) + \sum_{x=m,n,h,y} \frac{1}{2} \sigma^2 \, \partial^2_{u_x u_x} (a_x \mu_t) - \partial_{u_x} (b_x \mu_t) \tag{2.10}$$

with initial condition μ_0.

b) *There is a constant $C(T) > 0$ depending on V_0^{max} , $T > 0$ and on the coefficients of system* (2.1)*, but not on N, such that*

$$\sup_{t\in[0,T]} \mathbb{E}\left(\mathcal{W}_2^2(\mu_t^N, \mu_t)\right) \leq C(T)N^{-2/5}. \tag{2.11}$$

c) *If additionally functions ρ_x and ζ_x are of class $C^2(\mathbb{R})$, (or of class $C^1(\mathbb{R})$ when $\sigma = 0$), $(\mu_t : t \geq 0)$ given in part a) is the unique weak solution of* (2.10) *with initial condition μ_0 which has supports bounded uniformly in time.*

Remark 2.6.

i) *We have not been able to prove uniqueness of weak (measure) solutions to* (2.10) *in full generality. However, the global weak solution of* (2.10) *given by Theorem 2.5 a) is uniquely determined.*

ii) *Classically (see [39], [49]), convergence in law of μ_t^N to μ_t for fixed $t \geq 0$ implies the asymptotic independence as $N \to \infty$ of any subfamily $(X_t^{(1)}, \ldots, X_t^{(k)})$ of fixed size $k \leq N$ of system* (2.1) *(the propagation of chaos property). Somewhat counterintuitively, this is not incompatible with part a) of Theorem 2.3, even when $\sigma^2 = 0$.*

iii) *Parts a) and b) of Theorem 2.5 also hold for general exchangeable μ_0-chaotic initial conditions $(X_0^{(1)}, \ldots, X_0^{(N)})$ (that is, such that μ_0^N converges in law to μ_0 on $\mathcal{P}_2(\mathbb{R} \times [0,1]^4)$), in which case one must add a term of the form $C\mathbb{E}\left(\mathcal{W}_2^2(\mu_0^N, \mu_0)\right)$ on the right hand side of* (2.11).

iv) *The first assertion in Theorem 2.5 would be standard if the coefficients in* (2.1) *were globally Lipschitz. Under the key Hypothesis 2.2 we will be able to reduce the proof to the Lipschitz case. Moreover, this assumption will allow us to take full advantage of the estimates for empirical measures of i.i.d. samples proved in Fournier and Guillin [20], from where the convergence rate* (2.11) *will stem.*

Equation (2.10) can be interpreted as the dynamical description of a system of infinitely many HH neurons in mean field interaction. Thanks to Theorem 2.5 and to the uniformity in N of the results in Theorem 2.3, we can now finally transfer our synchronization results to this infinite dimensional setting. For each $t_1 \geq 0$, define $(\widehat{X}_t^{t_1,\infty} : t \geq t_1) = \left((\widehat{V}_t^\infty, \widehat{m}_t^\infty, \widehat{n}_t^\infty, \widehat{h}_t^\infty, \widehat{y}_t^\infty) : t \geq t_1\right)$ as the solution of the ordinary differential equation

$$\begin{aligned} \widehat{V}_t^\infty &= \widehat{V}_{t_1}^\infty + \int_{t_1}^t F(\widehat{V}_s^\infty, \widehat{m}_s^\infty, \widehat{n}_s^\infty, \widehat{h}_s^\infty) - J_{\text{Ch}}\widehat{y}_s^\infty(\widehat{V}_s^\infty - V_{\text{rev}})ds, \\ \widehat{x}_t^\infty &= \widehat{x}_{t_1}^\infty + \int_{t_1}^t \rho_x(\widehat{V}_s^\infty)(1-\widehat{x}_s^\infty) - \zeta_x(\widehat{V}_s^\infty)\widehat{x}_s^\infty ds, \quad x = m, n, h, y\,, \end{aligned} \tag{2.12}$$

with deterministic initial condition

$$\widehat{X}_{t_1}^{t_1,\infty} = \langle \mu_{t_1} \rangle,$$

where $(\mu_t : t \geq 0) \in C([0,\infty), \mathcal{P}_2(\mathbb{R} \times [0,1]^4))$ is the global weak solution of (2.10) with initial condition μ_0 given by Theorem 2.5 a). We have:

Corollary 2.7. *Under the assumptions of Theorem 2.5 and for the same constants as in Theorem 2.3, whenever $J_E > J_E^0$ we have:*

a) *For every $t \geq t_0$,*

$$\mathcal{W}_2^2(\mu_t, \delta_{\langle \mu_t \rangle}) \leq \mathcal{W}_2^2(\mu_{t_0}, \delta_{\langle \mu_{t_0} \rangle})e^{-\lambda^0(t-t_0)} + \sigma^2 \frac{C_{\zeta,\rho}^0}{\lambda^0}. \tag{2.13}$$

In particular, $\limsup_{t\to\infty} \mathcal{W}_2^2(\mu_t, \delta_{\langle \mu_t \rangle}) \leq \sigma^2 \frac{C_{\zeta,\rho}^0}{\lambda^0}$.

b) *For every $t_1 \geq t_0$ and $\delta \geq 0$ we have:*

$$\sup_{t_1 \leq t \leq t_1+\delta} \mathcal{W}_2^2(\mu_t, \delta_{\widehat{X}_t^{t_1,\infty}}) \leq K_0 \wedge 2\left[\left(K_0' e^{-\lambda_0(t_1-t_0)} + \sigma^2 \frac{C_{\zeta,\rho}^0}{\lambda^0}\right)(1+\delta K_\delta)\right].$$

We next present some numerical simulations which illustrate the validity of our theoretical results (at least from a qualitative point of view) and moreover we explore the behavior of system (2.1) when several neurons subpopulations are considered and when only chemical interaction is present. Furthermore, in view of the numerical experiments, we discuss some of the limitations and possible extensions of our theoretical results.

3 Numerical Experiments

Inspired in Bossy et al. [9], we have implemented numerical simulations of system (2.1) by means of an Exponential Projective Euler Scheme (EPES) which we next describe.

For a given time horizon $T > 0$ and a natural number M, we consider the time grid $\{t_0 = 0, t_1 = T/M, t_2 = 2T/M, \dots, t_k = kT/M, \dots, t_M = T\}$. As initial condition for each neuron in the system we consider independent random variables, uniformly distributed on $[-100, 100] \times [0, 1]^4$. Given the value of the system at t_k, the value for $\widehat{V}^{(i)}_{t_{k+1}}$ is computed as the exact solution to the ODE

$$\widehat{V}^{(i)}_t = \widehat{V}^{(i)}_{t_k} + \int_{t_k}^t F(\widehat{V}^{(i)}_s, \widehat{m}^{(i)}_{t_k}, \widehat{n}^{(i)}_{t_k}, \widehat{h}^{(i)}_{t_k})ds$$
$$- \int_0^t \frac{1}{N}\sum_{j=1}^N J_{\mathrm{E}}(\widehat{V}^{(i)}_s - \widehat{V}^{(j)}_{t_k}) - \frac{1}{N}\sum_{j=1}^N J_{\mathrm{Ch}}\widehat{y}^{(j)}_{t_k}(\widehat{V}^{(i)}_s - V_{\mathrm{rev}})ds.$$

which is indeed a linear ODE since F is linear in V. To compute $\widehat{x}^{(i)}_{t_{k+1}}$ we first solve the SDE

$$\breve{x}^{(i)}_t = \widehat{x}^{(i)}_{t_k} + \int_{t_k}^t \rho_x(\widehat{V}^{(i)}_{t_k})(1 - \breve{x}^{(i)}_s) - \zeta_x(\widehat{V}^{(i)}_{t_k})\breve{x}^{(i)}_s ds$$
$$+ \int_{t_k}^t \sigma_x(\widehat{V}^{(i)}_{t_k}, \widehat{x}^{(i)}_{t_k})dW^{x,i}_s, \quad x = m, n, h, y,$$

which corresponds to an Ornstein-Uhlenbeck process, so that $\breve{x}^{(i)}_{t_{k+1}}$ can be exactly simulated. However, since conditionally on $\widehat{x}^{(i)}_{t_j}, j \le k$ the law of $\breve{x}^{(i)}_{t_{k+1}}$ is Gaussian, $\{\breve{x}^{(i)}_{t_{k+1}} \notin [0, 1]\}$ happens with positive probability, so we are led to define $\widehat{x}^{(i)}_{t_{k+1}}$ by projecting $\breve{x}^{(i)}_{t_{k+1}}$ onto $[0, 1]$, that is:

$$\widehat{x}^{(i)}_{t_{k+1}} = \begin{cases} 0, & \breve{x}^{(i)}_{t_{k+1}} \in (-\infty, 0) \\ \breve{x}^{(i)}_{t_{k+1}}, & \breve{x}^{(i)}_{t_{k+1}} \in [0, 1] \\ 1, & \breve{x}^{(i)}_{t_{k+1}} \in (1, +\infty). \end{cases}$$

In Appendix E we prove the convergence in L^2-norm of the EPES applied to (2.1). We also provide the rate of convergence which is $1/2$ as for the classical Euler scheme.

In our simulations we have used as cut-off function (see Hypothesis 2.1-2)

$$\chi(u) = \begin{cases} 0.1 \exp\left(\frac{-0.5}{1-(2u-1)^2}\right) & u \in (0, 1) \\ 0 & u \notin (0, 1), \end{cases}$$

whereas, the specific values of the constants we have used are given in Table 1, and the rate functions ρ_x and ζ_x are given in Table 2 and shown in Figure 2. Although our results hold irrespectively of the value of input current I, we have taken in all simulations $I = 25$, in which case a noiseless single neuron with the chosen parameters has a limiting regime of sustained oscillations, see Figure 3.

g_{Na}	120 [mS/cm^3]	g_{K}	36 [mS/cm^3]	g_{L}	0.3 [mS/cm^3]
V_{Na}	50 [mV]	V_{K}	-77 [mV]	V_{L}	-54.4 [mV]

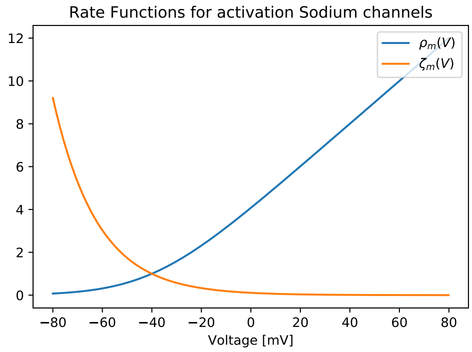
Table 1: Values for the constants in the function for F. Taken from [17, p.23]

3.1 Numerical experiments illustrating our theoretical results

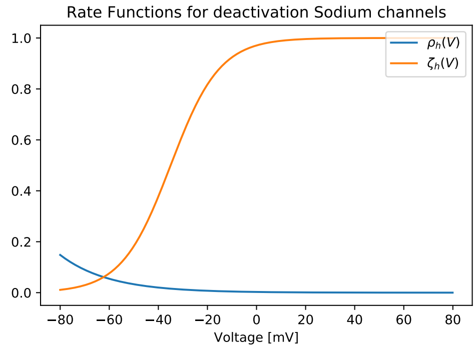
Our first numerical experiments illustrate the results of part a) in Theorem 2.3. In Figure 4 we show one trajectory of the system (2.1) under purely electrical interaction, for different sizes of the network and levels of noise. The first row shows the trajectories of a network of 10 neurons for $\sigma = 0$, $\sigma = 0.5$ and $\sigma = 1$. From the second to the fourth row, the trajectories of networks of 100, 1000 and 10000 neurons, respectively, are shown. The scale of all plots is the same. We observe that the qualitative behavior in terms of σ is the same for all rows: as expected, the noiseless network ultimately reaches perfect synchronization, whereas for $\sigma > 0$ the trajectories of the neurons lie in a band whose width

Channel type	$\rho_x(V)$	$\zeta_x(V)$
Sodium (Na) Activation Channels m	$\dfrac{0.1(V+40)}{1-\exp\left(-\dfrac{V+40}{10}\right)}$	$4\exp\left(-\dfrac{V+65}{18}\right)$
Sodium (Na) Deactivation Channels h	$0.07\exp\left(-\dfrac{V+65}{20}\right)$	$\dfrac{1}{1+\exp\left(-\dfrac{V+35}{10}\right)}$
Potassium (K) Activation Channels n	$\dfrac{0.01(V+55)}{1-\exp\left(-\dfrac{V+55}{10}\right)}$	$0.125\exp\left(-\dfrac{V+65}{80}\right)$
Neurotransmitter Channels y	$\dfrac{5}{1+\exp\left(-0.2(V-2.0)\right)}$	0.18

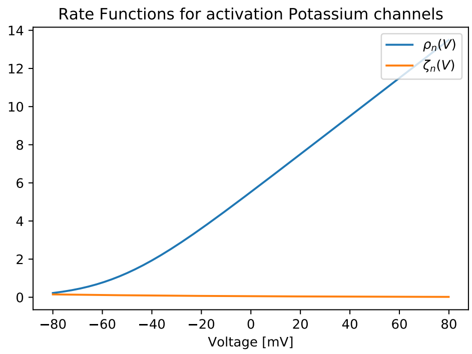
Table 2: Rate functions for the dynamics of the channels. Taken from [17, p.23] for the Sodium and Potassium channels, and from page [17, pp.160,163] for the neurotransmitter channel.



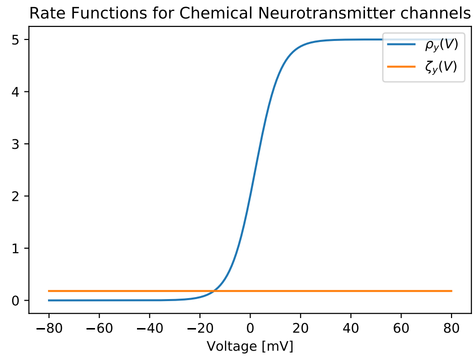
(a) Activation Sodium Channels.



(b) Deactivation Sodium Channels.



(c) Activation Potassium Channels.



(d) Synaptic Channels.

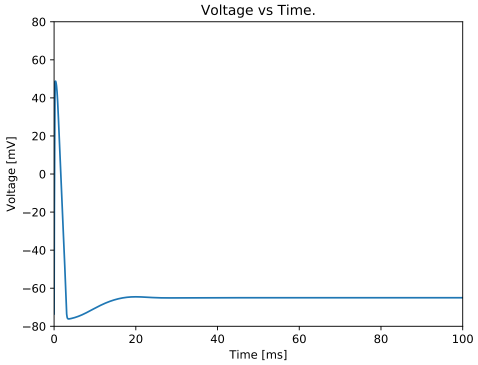
Figure 2: Charateristic plot of rate functions ρ_x and ζ_x.

increases with σ, as predicted by Theorem 2.3, a). Moreover, the speed at which synchronizations takes place does not depend on the size of the network nor on the level of noise.

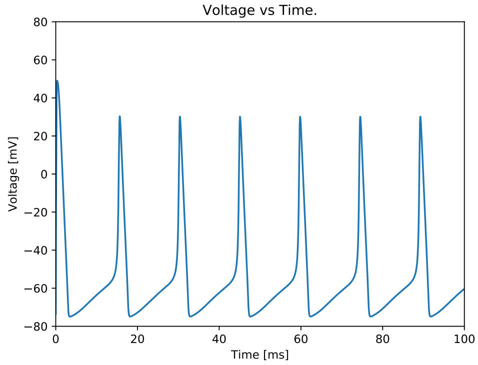
In our second experiment, we estimate the expected value of the empirical variance of a network of various sizes and for different levels of noise. More precisely we estimate the mean of

$$\bar{S}_t^V = \frac{1}{N}\sum_{i=1}^{N}\left(V_t^{(i)} - \bar{V}_t^N\right)^2, \quad \bar{S}_t^x = \frac{1}{N}\sum_{i=1}^{N}\left(x_t^{(i)} - \bar{x}_t^N\right)^2, \quad x = m, n, h.$$

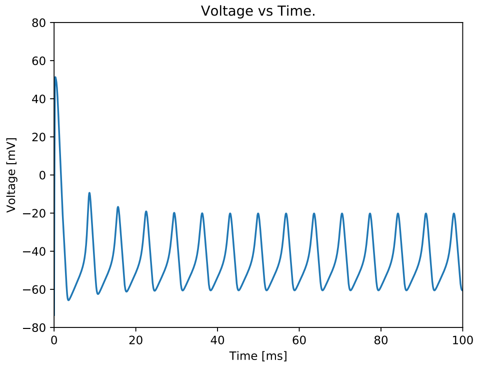
over 50000 Monte Carlo replica for each value of $\sigma \in \{0.1, 0.5, 1\}$ and $N \in \{10, 100, 1000, 10000\}$ (we now use $\sigma = 0.1$ instead of $\sigma = 0$ since in the latter case the obtained plot quickly becomes flat). The computation of the empirical variance for each time step and replica was done using the corrected two-phase algorithm to avoid *catastrophic cancellations* (see [14]). The results of this experiment appear in Figure 5, where a different variable is presented in each row, from top to bottom: voltage



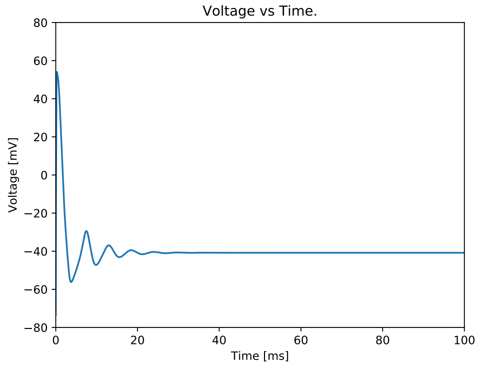
(a) $I = 0$.



(b) $I = 10$.



(c) $I = 100$.

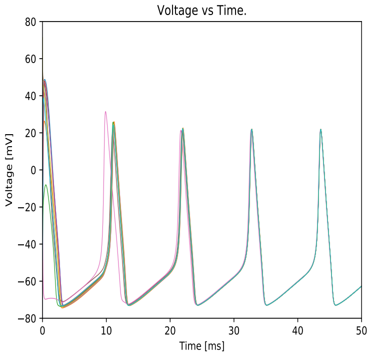


(d) $I = 200$.

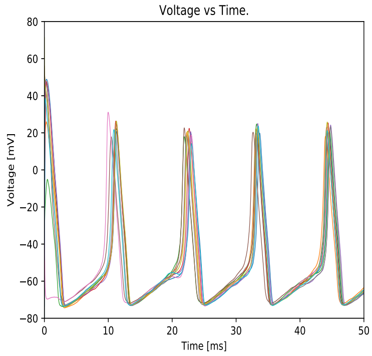
Figure 3: Reponses of the model (1.1) depending on the input current I: no oscillations if $I = 0$ (a); large amplitude and low frequency oscillations if $I = 10$ (b); small amplitude and high frequency oscillations if $I = 100$ (c); damped oscillations if $I = 200$ (d).

(V), Sodium activation channels (m), Potassium channels (n) and Sodium deactivation channels (h). Each column corresponds to a different level of noise, increasing from $\sigma = 0.1$ on the left, to $\sigma = 0.5$ in the middle and to $\sigma = 1$ on the right. In each subfigure we show the dissipation of the expected value of the empirical variance for networks of 10, 100, 1000 and 10000 neurons. Just as for one trajectory of the system, we observe again a quick synchronization, now measured in terms of the average dispersion over many trajectories, at speed which does not depend on the noise or the size of the network, with the heights of the peaks increasing with σ. Notice that double peaks are expected from Figure 5 already: even a small dispersion of the phase among different neurons can induce a high dispersion of their voltages and channels right before and after a potential spike is emitted. This dispersion increases with N, but tends to stabilize as N becomes large (notice that the red and green lines in Figure 5 are indistinguishable), consistently with Theorem 2.5.

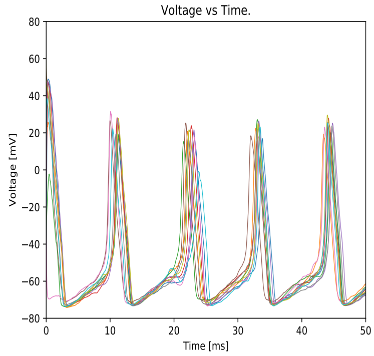
From this last observation, it is also interesting to point out that the maximum variance over time-windows of fixed length $\delta > 0$ which drift to infinity cannot decrease for every possible value of δ, unless $\sigma^2 = 0$. Indeed, in the noisy case the voltage of significantly many neurons can in principle differ from the voltage of the underlying one-neuron dynamics, over time-windows larger than its period, by as much as the whole asymptotic range of the voltages dynamics. Thus, albeit not sharp, the estimates in part b) of Theorem 2.3 and Corollary 2.7 are qualitatively correct.



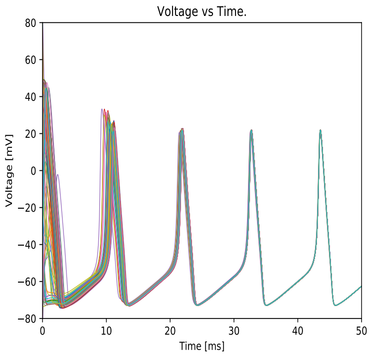
(a) $N = 10$, $\sigma = 0$.



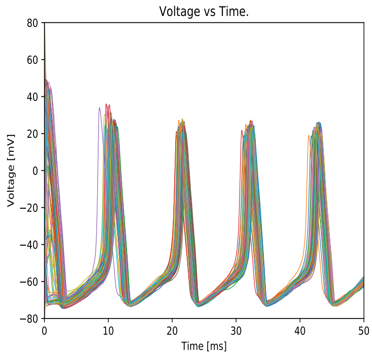
(b) $N = 10$, $\sigma = 0.5$.



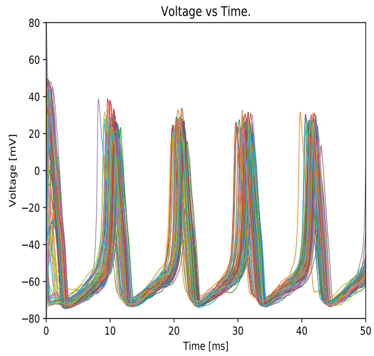
(c) $N = 10$, $\sigma = 1$.



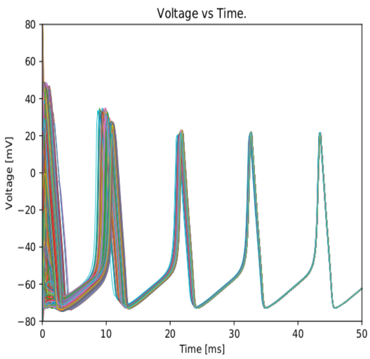
(d) $N = 100$, $\sigma = 0$.



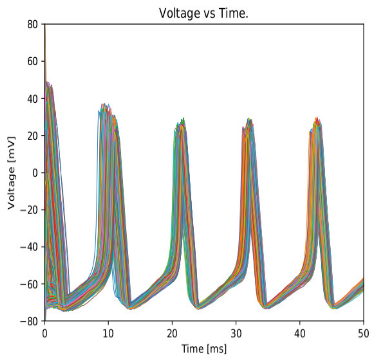
(e) $N = 100$, $\sigma = 0.5$.



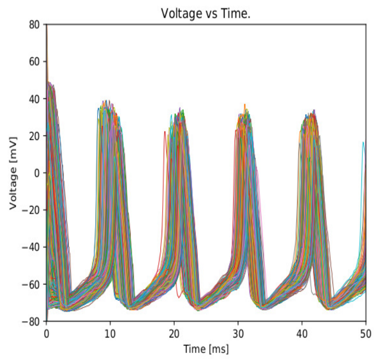
(f) $N = 100$, $\sigma = 1$.



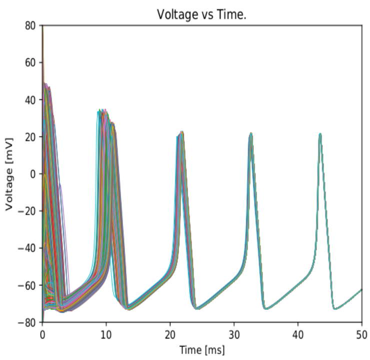
(g) $N = 1000$, $\sigma = 0$.



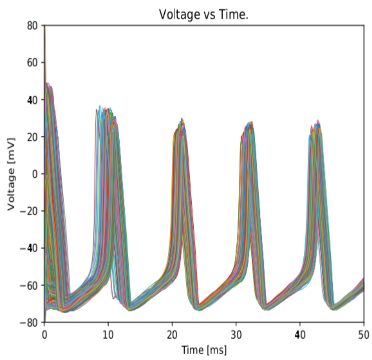
(h) $N = 1000$, $\sigma = 0.5$.



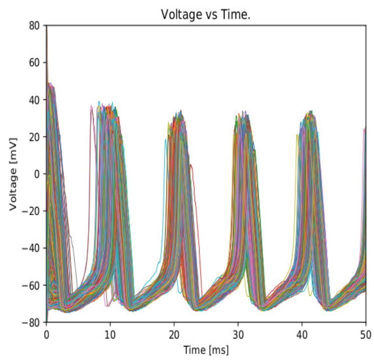
(i) $N = 1000$, $\sigma = 1$.



(j) $N = 10000$, $\sigma = 0$.

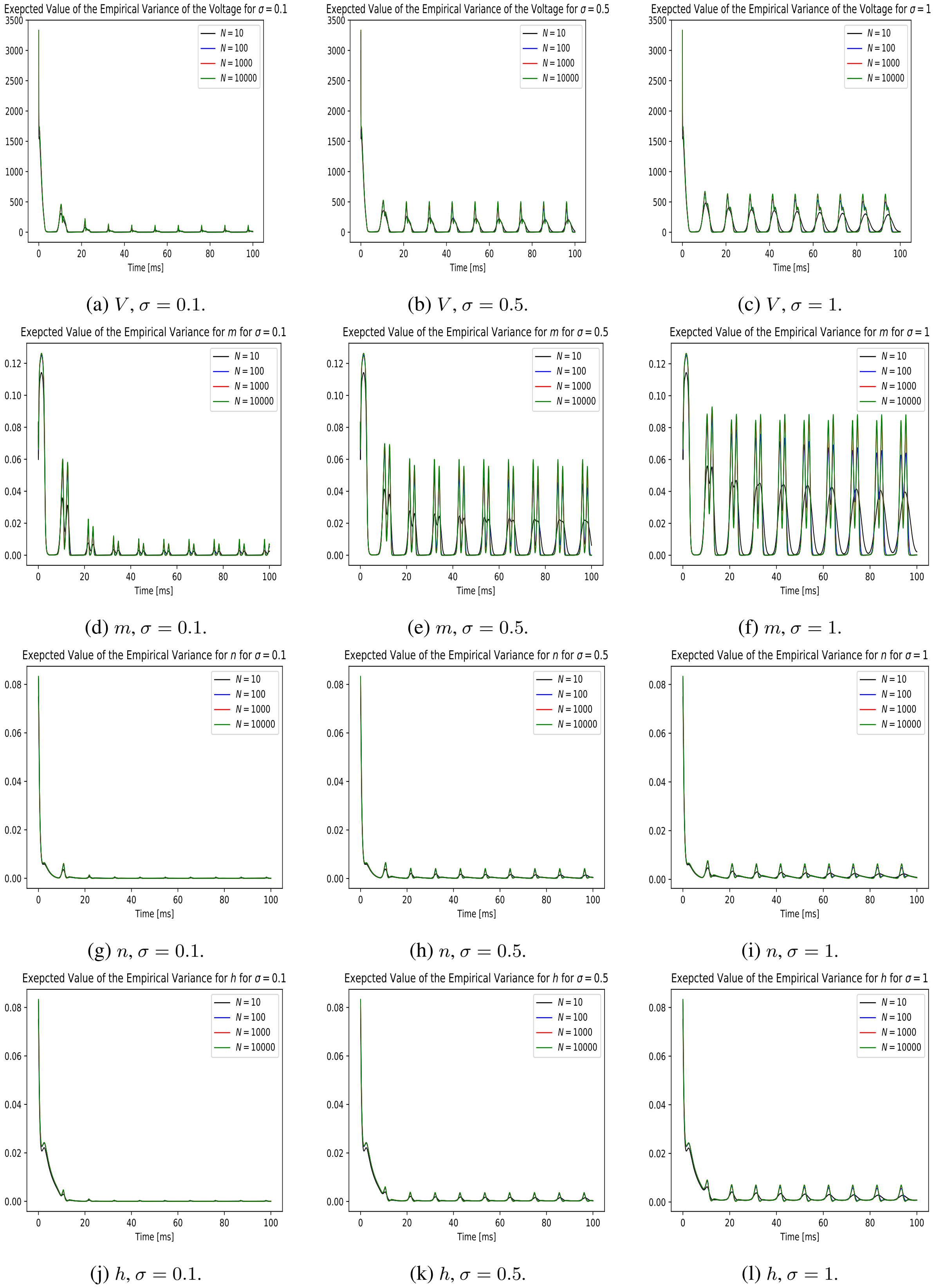


(k) $N = 10000$, $\sigma = 0.5$.



(l) $N = 10000$, $\sigma = 1$.

Figure 4: Synchronization of a network under pure electrical interaction for different network sizes N and noise levels σ.



(a) V, $\sigma = 0.1$. (b) V, $\sigma = 0.5$. (c) V, $\sigma = 1$.

(d) m, $\sigma = 0.1$. (e) m, $\sigma = 0.5$. (f) m, $\sigma = 1$.

(g) n, $\sigma = 0.1$. (h) n, $\sigma = 0.5$. (i) n, $\sigma = 1$.

(j) h, $\sigma = 0.1$. (k) h, $\sigma = 0.5$. (l) h, $\sigma = 1$.

Figure 5: Dissipation of the empirical variance for different level of noise. First row: expected empirical variance of the voltage; second to fourth rows: expected empirical variance of the Sodium activation channels (m), the Potassium channels (n) and the Sodium deactivation channels (h) respectively. $J_{\text{Ch}} = 0$ and $J_{\text{E}} = 1$ in all simulations.

3.2 Beyond Theorem 2.3

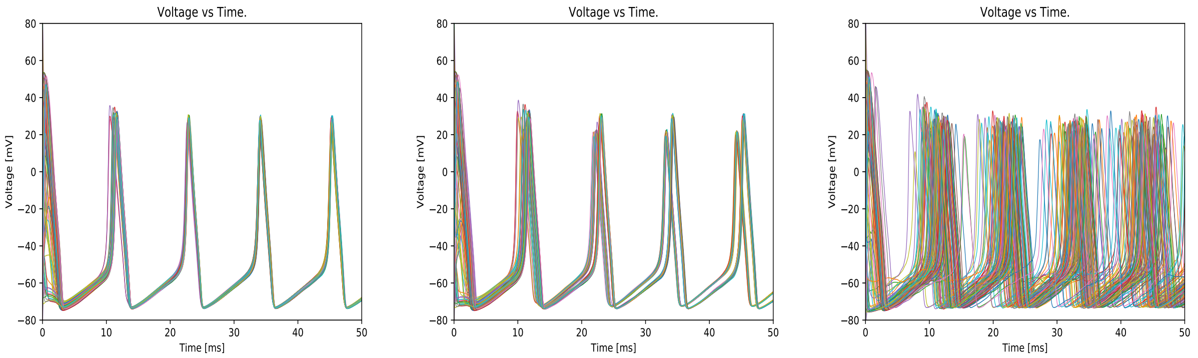
We next carry out two type of experiments in situations not covered by our theoretical results on synchronization. In the first one we study the behavior of a more realistic network with several subpopulations of neurons. The dynamics of the i-th neuron in the subpopulation of type $\alpha \in P$, with P denoting the set of subpopulations, is given by

$$
\begin{aligned}
V_t^{(i)} &= V_0^{(i)} + \int_0^t F_\alpha(V_s^{(i)}, m_s^{(i)}, n_s^{(i)}, h_s^{(i)})ds \\
&\qquad - \int_0^t \sum_{\gamma \in P} \frac{1}{N_\gamma} \sum_{j=1}^{N_\gamma} J_{\mathrm{E}}^{\alpha\gamma}(V_s^{(i)} - V_s^{(j)}) - \sum_{\gamma \in P} \frac{1}{N_\gamma} \sum_{j=1}^{N_\gamma} J_{\mathrm{Ch}}^{\alpha\gamma} y_s^{(j)}(V_s^{(i)} - V_{\mathrm{rev}}^{\alpha\gamma})ds, \\
x_t^{(i)} &= x_0^{(i)} + \int_0^t \rho_x^\alpha(V_s^{(i)})(1 - x_s^{(i)}) - \zeta_x^\alpha(V_s^{(i)})x_s^{(i)}ds \\
&\qquad\qquad + \int_0^t \sigma_x^\alpha(V_s^{(i)}, x_s^{(i)})dW_s^{x,i}, \quad x = m, n, h, y\,, \qquad t \geq 0\,,
\end{aligned}
$$

where N_γ is the number of neurons in subpopulation γ. We notice that in this case the electric and chemical conductivity parameters are $|P| \times |P|$ matrices. Propagation of chaos for such systems as $N \to \infty$ was proved in [10] (though under slightly more stringent assumptions on the coefficients). In Figure 6 we show one trajectory of a network of 100 neurons with two subpopulations, each of them with 50 neurons. On the left (plot (a)), we consider the two subpopulations with different levels of noise and different input current for each of them (different F), meanwhile the electrical conductance matrix $J_{\mathrm{E}}^{\alpha\gamma}$ is homogeneous, with all the components equal to 1. We observe that the whole network gets synchronized. We believe that Theorem 2.3 can be easily extended to this case (or, more generally, when $\inf_{\alpha,\gamma \in P} J_{\mathrm{E}}^{\alpha\gamma}$ is big enough). In the middle (plot (b)), we observe that, if in addition to considering different subpopulations, the matrix J_{E} is not homogeneous (taking in some entries strictly smaller values than the largest value 1), then synchronization can be observed in each subpopulation but not globally. More precisely, in these examples the two populations synchronize out-of-phase. Finally, on the right (plot (c)), we observe no evidence of synchronization when neurons in one population with same F are electrically connected with two small but different values for J_{E}. This is in line with our theoretical result that thresholds the synchronization of the dynamics for a big enough $\inf_{\alpha,\gamma \in P} J_{\mathrm{E}}^{\alpha\gamma}$, even if Theorem 2.3 gives only a sufficient condition on J_{E}.

According to Kopell and Ermentrout [34], “a small amount of electrical conductance can increase the degree of synchronization far more that a much larger increase in inhibitory conductance”. This is consistent with what we have observed in our numerical experiments. The effect of the electrical interaction is stronger and faster than the effect of the chemical interaction. Therefore to appreciate the effect of the chemical synapses, the second type of experiment we have performed concerns only chemical synapses, that is, $J_{\mathrm{E}} = 0$. In Figure 7 we show one trajectory of a network of 100 neurons interacting through inhibitory chemical synapses (in this case, we choose $V_{\mathrm{rev}} = -75$ according to [17, p.163]). We observe an anti-phase synchronization that persists in time (see in plot (a) the transition to the stationary regime in plot (b)), in which two clusters of simultaneously firing neurons emerge. We note that the relative sizes of the two clusters is random and might change from one simulation to another one.

On the other hand, Figure 8 shows one trajectory of a network of 100 neurons interacting through excitatory chemical synapses (with $V_{\mathrm{rev}} = 0$, see [17, p.161]). Some kind of synchronization, similar to the case of purely electrical synapses (see e.g. Figure 6(a)), emerges also here, although the shape of the oscillations is different and the frequency of the spikes is smaller.

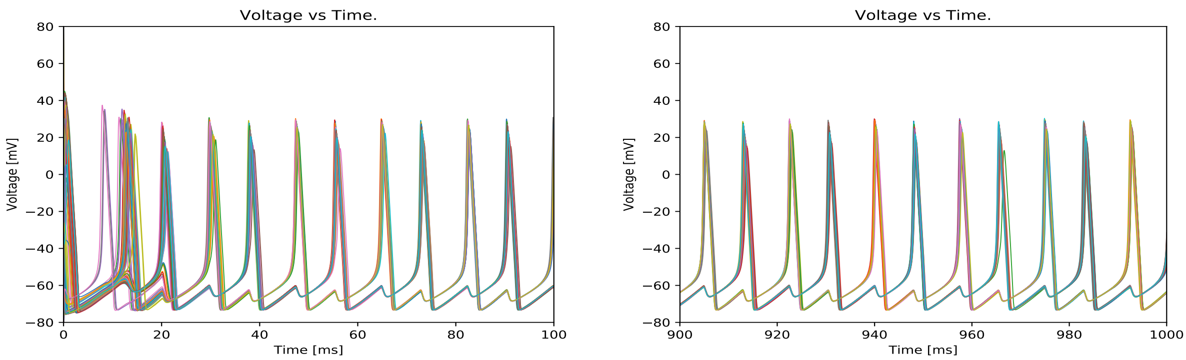


(a) Populations are different, the interaction is homogeneous.

(b) Populations are different, the interaction is heterogeneous.

(c) Populations are equal, the interaction is heterogeneous.

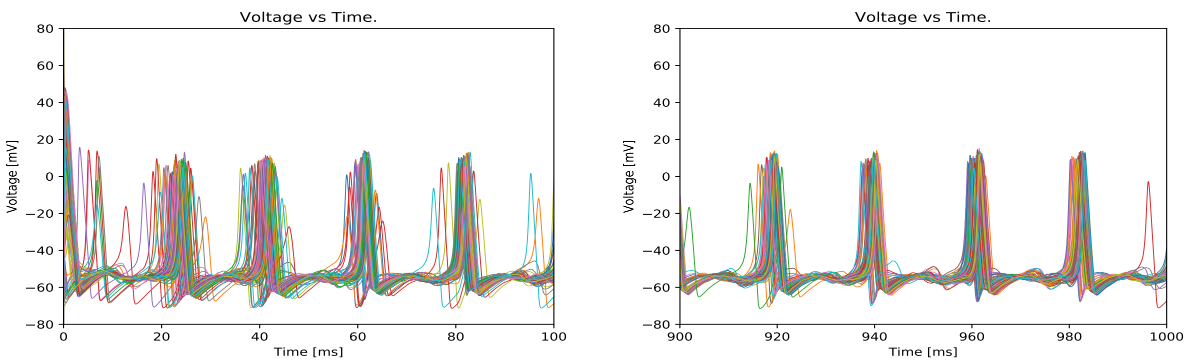
Figure 6: Trajectories for network of 100 neuron with two subpopulations.



(a) Voltage of 100 neurons at the beginning of the simulation.

(b) Voltage of 100 neurons after 900 [ms].

Figure 7: Trajectories for network of 100 neuron with inhibitory chemical synapses



(a) Voltage of 100 neurons at the beginning of the simulation.

(b) Voltage of 100 neurons after 900 [ms].

Figure 8: Trajectory of a network of 100 neurons with excitatory chemical synapses

4 Concluding remarks and open problems

The numerical experiments presented in Section 3.1 show that, from a quantitative point of view, our theoretical results should still allow for considerable improvements. Indeed, our simulations indicate that the actual global bounds on the voltage, the critical interaction strength above which synchronization happens and the asymptotic discrepancy from synchronization are much smaller than suggested from rough estimates of the bounds in our theoretical results and their proofs. In turn, the actual exponential synchronization rate seems to be much higher. Also, our theoretical results treat the values of the voltage and channels variables jointly, although they are of considerably different orders of magnitude (i.e. channel variables and variances are negligible compared to the voltages ones). The numerical experiments show in turn that the theoretically described behavior happens at each variable's scale.

We must also emphasize that the anti- or out-of-phase- synchronization responses put in evidence in the numerical experiments presented in Section 3.2 are not well captured by the empirical variance criteria proposed in our Theorem 2.3. In those cases, a natural, though challenging strategy would be to extend the phase reduction approach and results developed e.g. in [28] and [27] for finite deterministic networks of HH neurons in order to obtain rigorous synchronization results, regardless of the network size, and then in the mean field limit. Another interesting but also challenging question is studying the existence of stationary measures for the McKean-Vlasov limit equation in relation with some characteristic behavior of the system and its possible synchronization regimes, in the vein of recent works of Bertini et al. [8], Giacomin et al. [24] and Luçon and Poquet [37]. These questions are left for future works.

A Basic properties of the model (2.1)

We start establishing three basic facts about the system of stochastic differential equations (2.1): its (strong) global well-posedness, the fact that the open channels proportion processes stay (as required) in $[0,1]$ and, finally, and explicit global bound for the voltage processes in terms of a bound for the initial values.

Lemma A.1. *Assume Hypothesis 2.1. Then, strong existence and pathwise uniqueness holds for system (2.1). Moreover, a.s. for all $t \geq 0$ and every $i = 1, \dots, N$ we have $(m_t^{(i)}, n_t^{(i)}, h_t^{(i)}, y_t^{(i)}) \in [0,1]^4$. In particular, the absolute value in (2.2) can be removed.*

Proof. It is enough to prove the result for deterministic initial data so we assume this is the case. Take $M > 0$ fixed, and for $j = 1, 3, 4$ define truncation functions p_M^j on $\mathbb{R}$ by

$$p_M^j(x) = \begin{cases} x^j & x \in [-M, M] \\ M^j & x \in (M, \infty) \\ (-M)^j & x \in (-\infty, -M). \end{cases}$$

Let $X^{(M)} := (X^{(1,M)}, \dots, X^{(N,M)})$ with $X_t^{(i,M)} = (V_t^{(i,M)}, m_t^{(i,M)}, n_t^{(i,M)}, h_t^{(i,M)}, y_t^{(i,M)})$, $i = 1, \dots, N$ be defined by

$$\begin{aligned} V_t^{(i,M)} &= V_0^{(i,M)} + \int_0^t F_M(V_s^{(i,M)}, m_s^{(i,M)}, n_s^{(i,M)}, h_s^{(i,M)}) - \frac{1}{N}\sum_{j=1}^N J_{\mathrm{E}}(V_s^{(i,M)} - V_s^{(j,M)}) \\ &\qquad - \frac{1}{N}\sum_{i=1}^N J_{\mathrm{Ch}} p_M^1(y_s^{(j,M)})(p_M^1(V_s^{(i,M)}) - V_{\mathrm{rev}})ds, \\ x_t^{(i,M)} &= x_0^{(i,M)} + \int_0^t \rho_x(p_M^1(V_s^{(i,M)}))(1 - p_M^1(x_s^{(i,M)})) - \zeta_x(p_M^1(V_s^{(i,M)}))p_M^1(x_s^{(i,M)})ds \\ &\qquad + \int_0^t \sigma_x(p_M^1(V_s^{(i,M)}), x_s^{(i,M)})dW_s^{x,i}, \quad x = m, n, h, y, \end{aligned} \tag{A.1}$$

where

$$F_M(v,m,n,h) = I - g_{\mathrm{K}} p_M^4(n)(p_M^1(v) - V_{\mathrm{K}}) - g_{\mathrm{Na}} p_M^3(m) p_M^1(h)(p_M^1(v) - V_{\mathrm{Na}}) - g_{\mathrm{L}}(v - V_{\mathrm{L}}). \tag{A.2}$$

Is is immediate that the drift coefficients in system (A.1) are Lipschitz continuous. This is less clear in the case of the diffusion coefficients, so we check this point next. Notice that

$$\max_{(v,u)\in\mathbb{R}\times[0,1]} \rho_x(p^1_M(v))(1-u)+\zeta_x(p^1_M(v))u \le S_M := \max_{v\in[-M,M]} \rho_x(v)+\zeta_x(v) < \infty$$

whereas, thanks to point 2) in Hypothesis 2.1,

$$\min_{(v,u)\in\mathbb{R}\times[0,1]} \rho_x(p^1_M(v))(1-u)+\zeta_x(p^1_M(v))u \ge \delta_M := \min_{v\in[-M,M]} \{\rho_x(v),\zeta_x(v)\} > 0.$$

Therefore, one can find a bounded Lipschitz continuous function $g_x : \mathbb{R} \to \mathbb{R}_+$ such that $g_x(s) = \sqrt{s}$ on $(\delta_M/2, 2S_M)$ and rewrite the diffusion coefficients in (A.1) as

$$\sigma_x(p^1_M(v),u) = \sigma g_x(|\rho_x(p^1_M(v))(1-u)+\zeta_x(p^1_M(v))u|)\chi(u).$$

It is then easily seen that $|\sigma_x(p^1_M(v),u)-\sigma_x(p^1_M(v'),u')| \le C_M(|u-u'|+|v-v'|)$ for some $C_M > 0$ in each of the three cases $(u,u') \in [0,1]^2$, $(u,u') \in ([0,1]^2)^c$ and $(u,u') \in [0,1]\times[0,1]^c$ for any $v,v' \in \mathbb{R}$. Thus, global pathwise well-posedness for system (A.1) holds.

Thanks to the second assumption in point 4) of Hypothesis 2.1 and the fact that $\sigma_x(v,u) = 0$ for $(v,u) \in \mathbb{R}\times(0,1)^c$ and $[\rho_x(v)(1-u)-\zeta_x(v)u][\mathbf{1}_{(-\infty,0]}(u)-\mathbf{1}_{[1,-\infty)}(u)] \ge 0$ for $(v,u) \in \mathbb{R}^2$, we can more apply Proposition 3.3 in [10] to get that $x^{(1,M)},\dots,x^{(N,M)}$ are confined in $[0,1]$ for all time (notice that the proof of that result still works if Hypothesis 2.1 i) therein that χ be compactly supported in $(0,1)$ is replaced by χ being supported in $[0,1]$).

We can now use standard arguments to deduce global existence and pathwise uniqueness of a solution to system (2.1). Indeed, setting $\theta_M = \inf\{t \ge 0 : |X^{(M)}_t| \ge M\}$, using the global Lipschitz character of its coefficients together with Itô calculus and Gronwall's lemma we get for every $M' > M$ that a.s. for all $t \ge 0$, $X^{(M)}_{t\wedge\theta_M} = X^{(M')}_{t\wedge\theta_M}$. This implies that $\theta_{M'} > \theta_M$ a.s. and allows us to unambiguously define a process X solving (2.1) on the random interval $[0,\theta)$, with $\theta := \sup_{M>0}\theta_M$, by $X_t = X^{(M)}_t$ for all $t \in [0,\theta_M]$. On the other hand, since $|p^1_M(z)| \le |z|$ for all $z \in \mathbb{R}$, for two constants $C_1, C_2 > 0$ not depending on $M > 0$ we have $|F_M(v,m,n,h)| \le C_1 + C_2|v|$ for every $(v,m,n,h) \in \mathbb{R}\times[0,1]^3$. Using this control on the right hand side of the equations for $V^{(1,M)},\dots,V^{(N,M)}$ in (A.1) and Gronwall's lemma we get

$$\mathbb{E}\left[|X_{t\wedge\theta_M}|\right] \le C(t),$$

for some constant $C(t) > 0$ not depending on M. This yields $M\mathbb{P}\left[\theta_M < t\right] \le C(t)$, whence $\mathbb{P}\left[\theta < \infty\right] = 0$ letting M and then $t \nearrow \infty$. The statement follows. □

Remark A.2. *i) The arguments given in the previous proof also show that each of the functions σ_x is locally Lipschitz on $\mathbb{R}\times[0,1]$.*

ii) The same proof also works for some extensions of our model. For instance, if independent Brownian motions are added to each of the voltage processes.

We next show that under the additional Hypothesis 2.2, each of the voltage processes is bounded uniformly in time and in N. Below and in all the sequel we denote

$$V^{\text{max}}_{t,\infty} := \max_{i=1,\dots,N}\sup_{s\in[t,\infty)} |V^{(i)}_s|.$$

We also set

$$R_{\text{max}} := \max_{r,s,u\in[0,1]} |I + g_{\text{Na}}V_{\text{Na}}r + g_{\text{K}}V_{\text{K}}s + g_{\text{L}}V_{\text{L}} + J_{\text{Ch}}V_{\text{rev}}u|.$$

Proposition A.3. *Under Hypothesis 2.2, for every $N \ge 1$ and $t \ge 0$ we have a.s.*

$$\left|\bar{V}^N_t\right| \le V^{max}_0 e^{-g_L t} + \frac{2R_{max}}{g_L}(1-e^{-g_L t})$$

and

$$V^{max}_{t,\infty} \le V^*_t := \frac{4R_{max}}{g_L} + 2V^{max}_0 e^{-g_L t}. \tag{A.3}$$

As a consequence for every $N \ge 1$, there exists at least one invariant law μ^N_∞ for the solution to (2.1), namely there exists a solution $(X_t, t \ge 0)$ to (2.1) such that X_t has law μ^N_∞ for all $t \ge 0$ as soon as X_0 has law μ^N_∞. Moreover, this invariant measure is exchangeable.

Remark A.4.

i) *The bound V_t^* on $V_{t,\infty}^{max}$ is in general not optimal. For instance, if $V_0^{max} < \frac{2R_{max}}{g_L}$, one can choose $V_0^{max} > \frac{2R_{max}}{g_L+J_E}$ and get from the last identity in (A.5) that $V_{0,\infty}^{max} \le \frac{4R_{max}}{g_L} < V_0^*$. However, in order to state a synchronization result that holds for a general class of initial conditions V_0, the fact that the bound V_t^* does not depend on the electrical connectivity J_E and that $V_\infty^* := \lim_{t\to\infty} V_t^* = \frac{4R_{max}}{g_L}$ does not depend on the initial condition will be crucial. See point i) in Remark B.4 for a related discussion.*

ii) *If point 2) of Hypothesis 2.2 does not hold, by slightly modifying the arguments of Lemma A.3 we still can get the a.s. bound*

$$\left|V_t^{(i)}\right| \le \frac{4R_{max}}{g_L} + 2\frac{|V_0|}{\sqrt{N}}e^{-g_L t}$$

implying a uniform in N bound for $\mathbb{E}(V_{t,\infty}^{max})$ if for instance all the random variables $V_0^{i,N}$, $i = 1,\dots,N$, $N \ge 1$ are equal in law and have finite second moment. However, we have not been able to fully extend our results to such a framework.

iii) *The same arguments also show that a bound like (A.3) holds with $V_{t,\infty}^{max}$ replaced by*

$$\widehat{V^{max}}_{t,\infty} := \max_{i=1,\dots,N} \sup_{s\in[t,\infty)} |\widehat{V}_s^{(i)}|.$$

That is, the voltages obtained with the EPE scheme are also uniformly bounded.

In the proof of Proposition A.3 and later, we will make use of the the following version of Gronwall's lemma (see e.g. Ambrosio et al. [1, p. 88]).

Lemma A.5. *Let $\theta : [0,+\infty) \to \mathbb{R}$ be a locally absolutely continuous function and $a, b \in L^1_{loc}([0,+\infty))$ be given functions satisfying, for $\lambda \in \mathbb{R}$,*

$$\frac{d}{dt}\theta^2(t) + 2\lambda\theta^2(t) \le a(t) + 2b(t)\theta(t) \text{ for } \mathcal{L}^\infty - a.e.\ t > 0.$$

Then for every $T > 0$ we have

$$e^{\lambda T}|\theta(T)| \le \left(\theta^2(0) + \sup_{t\in[0,T]}\int_0^t e^{2\lambda s}a(s)ds\right)^{1/2} + 2\int_0^T e^{\lambda t}|b(t)|dt.$$

Proof of Proposition A.3. Setting

$$R_s^{(i)} := I + g_{\text{Na}}V_{\text{Na}}\left[m_s^{(i)}\right]^3 h_s^{(i)} + g_{\text{K}}V_{\text{K}}\left[n_s^{(i)}\right]^4 + g_{\text{L}}V_{\text{L}} + J_{\text{Ch}}V_{\text{rev}}\bar{y}_s^N, \text{ and}$$

$$A_s^{(i)} := g_{\text{Na}}\left[m_s^{(i)}\right]^3 h_s^{(i)} + g_{\text{K}}\left[n_s^{(i)}\right]^4 + g_{\text{L}} + J_{\text{Ch}}\bar{y}_s^N,$$

the dynamics of the potential can be written as

$$V_t^{(i)} = V_0^{(i)} + \int_0^t R_s^{(i)} - A_s^{(i)}V_s^{(i)} - J_{\text{E}}V_s^{(i)} + J_{\text{E}}\bar{V}_s^N ds.$$

Therefore, we get

$$\left(V_t^{(i)}\right)^2 = \left(V_0^{(i)}\right)^2 + 2\int_0^t R_s^{(i)}V_s^{(i)} - A_s^{(i)}\left(V_s^{(i)}\right)^2 - J_{\text{E}}\left(V_s^{(i)}\right)^2 + J_{\text{E}}\bar{V}_s^N V_s^{(i)} ds. \tag{A.4}$$

and

$$|V_t|^2 = |V_0|^2 + 2\int_0^t \sum_{i=1}^N \left[R_s^{(i)}V_s^{(i)} - A_s^{(i)}\left(V_s^{(i)}\right)^2\right] - J_{\text{E}}|V_s|^2 + NJ_{\text{E}}(\bar{V}_s^N)^2 ds.$$

Notice that

$$(\bar{V}_s^N)^2 = \frac{1}{N^2}\sum_{i,j=1}^N V_s^{(i)}V_s^{(j)} \le \frac{1}{2N^2}\sum_{i,j=1}^N (V_s^{(i)})^2 + (V_s^{(j)})^2 = \frac{1}{N}|V_s|^2,$$

which yields

$$\frac{d}{dt}|V_t|^2 + 2g_{\mathrm{L}}|V_t|^2 \le 2|R_t||V_t|.$$

By Lemma A.5 we deduce that

$$|V_t| \le |V_0|e^{-g_{\mathrm{L}}t} + 2e^{-g_{\mathrm{L}}t}\int_0^t e^{g_{\mathrm{L}}s}|R_s|ds.$$

Since $|R_s^{(i)}| \le R_{\max}$, we then get

$$\left|\bar{V}_t^N\right| \le \frac{|V_t|}{\sqrt{N}} \le \frac{|V_0|}{\sqrt{N}}e^{-g_{\mathrm{L}}t} + \frac{2R_{\max}}{g_{\mathrm{L}}}(1-e^{-g_{\mathrm{L}}t}) \le V_0^{\max}e^{-g_{\mathrm{L}}t} + \frac{2R_{\max}}{g_{\mathrm{L}}}(1-e^{-g_{\mathrm{L}}t}).$$

which is the first desired inequality. Using this in (A.4) yields

$$\begin{aligned}\frac{d}{dt}(V_t^{(i)})^2 + 2(g_{\mathrm{L}}+J_{\mathrm{E}})(V_t^{(i)})^2 &\le 2|R_t^{(i)} + J_{\mathrm{E}}\bar{V}_t^N||V_t^{(i)}| \\ &\le 2\left(R_{\max} + J_{\mathrm{E}}\left(V_0^{\max} - \frac{2R_{\max}}{g_{\mathrm{L}}}\right)e^{-g_{\mathrm{L}}t} + \frac{2J_{\mathrm{E}}R_{\max}}{g_{\mathrm{L}}}\right)|V_t^{(i)}|.\end{aligned}$$

Applying once again Lemma A.5, we obtain

$$\begin{aligned}\left|V_t^{(i)}\right| &\le V_0^{\max}e^{-(g_{\mathrm{L}}+J_{\mathrm{E}})t} \\ &\quad + 2e^{-(g_{\mathrm{L}}+J_{\mathrm{E}})t}\int_0^t e^{(g_{\mathrm{L}}+J_{\mathrm{E}})s}\left(R_{\max}\left(\frac{g_{\mathrm{L}}+2J_{\mathrm{E}}}{g_{\mathrm{L}}}\right) + J_{\mathrm{E}}\left(V_0^{\max} - \frac{2R_{\max}}{g_{\mathrm{L}}}\right)e^{-g_{\mathrm{L}}s}\right)ds \\ &= V_0^{\max}e^{-(g_{\mathrm{L}}+J_{\mathrm{E}})t} + 2R_{\max}\left(\frac{g_{\mathrm{L}}+2J_{\mathrm{E}}}{g_{\mathrm{L}}(g_{\mathrm{L}}+J_{\mathrm{E}})}\right)(1-e^{-(g_{\mathrm{L}}+J_{\mathrm{E}})t}) \\ &\quad + 2\left(V_0^{\max} - \frac{2R_{\max}}{g_{\mathrm{L}}}\right)\left(e^{-g_{\mathrm{L}}t} - e^{-(g_{\mathrm{L}}+J_{\mathrm{E}})t}\right) \\ &= \frac{2R_{\max}}{g_{\mathrm{L}}}\left(\frac{g_{\mathrm{L}}+2J_{\mathrm{E}}}{g_{\mathrm{L}}+J_{\mathrm{E}}}\right) + 2\left(V_0^{\max} - \frac{2R_{\max}}{g_{\mathrm{L}}}\right)e^{-g_{\mathrm{L}}t} \\ &\quad + \left(\frac{2R_{\max}}{g_{\mathrm{L}}+J_{\mathrm{E}}} - V_0^{\max}\right)e^{-(g_{\mathrm{L}}+J_{\mathrm{E}})t} \\ &\le \frac{4R_{\max}}{g_{\mathrm{L}}} + 2V_0^{\max}e^{-g_{\mathrm{L}}t} = V_t^* \end{aligned} \tag{A.5}$$

which implies the asserted bounds on $V_{t,\infty}^{\max}$.

Let us now deduce the existence of an invariant distribution which is exchangeable. Let P_t^N denote the semigroup associated to the solution of (2.1), that is for each $\mathcal{X} \in (\mathbb{R}\times[0,1]^4)^N$ and B Borel set of $(\mathbb{R}\times[0,1]^4)^N$,

$$P_t^N(\mathcal{X}, B) = \mathbb{P}\left(X_t \in B | X_0 = \mathcal{X}\right).$$

Consider also the probability measure $R_T^N(\lambda)$ on $(\mathbb{R}\times[0,1]^4)^N$, defined for any law λ as

$$R_T^N(\lambda)(B) = \int_{(\mathbb{R}\times[0,1]^4)^N}\left(\frac{1}{T}\int_0^T P_t^N(\mathcal{X}, B)dt\right)\lambda(d\mathcal{X}).$$

Since the voltage component is uniformly bounded in time, by (A.5), the solution to (2.1) lies in the compact set $([-4\frac{R_{\max}}{g_{\mathrm{L}}} - 2V_0^{\max}, 4\frac{R_{\max}}{g_{\mathrm{L}}} + 2V_0^{\max}]\times[0,1]^4)^N$, and then for any $(T_M) \nearrow \infty$, and any λ with compact support, the sequence $(R_{T_M}^N(\lambda), M \ge 0)$ is tight and has a subsequence weakly converging to some probability measure μ_∞^N. According to Krylov-Bogoliubov Theorem, μ_∞^N is invariant for P_t^N.

Let us now choose and exchangeable initial law λ. For any measurable and bounded function ψ, the identity

$$\int_{(\mathbb{R}\times[0,1]^4)^N} P_t^N(\mathcal{X}, dy)\psi(y)\lambda(d\mathcal{X}) = \int_{(\mathbb{R}\times[0,1]^4)^N} P_t^N(\mathcal{X}, dy)(\psi\circ\pi)(y)\lambda(d\mathcal{X}).$$

for any N-permutation π of the coordinates follows directly from the exchangeable structure of the system of equations (2.1). Therefore, $R_{T_M}^N(\lambda)$ is exchangeable for any T_M, and the corresponding μ_∞^N is exchangeable as the weak limit of exchangeable measures.

□

B Synchronization: proof of Theorem 2.3 *a)*

In the sequel, for any locally bounded real function f on $\mathbb{R}$ and each $R > 0$ we will write

$$\|f\|_{\infty,R} := \sup_{v\in[-R,R]} |f(v)|.$$

We will repeatedly use a simple control of the increments of the function F, stated in next lemma for convenience:

Lemma B.1. *We have*

$$\begin{aligned}(F(V_1,m_1,n_1,h_1) - F(V_2,m_2,n_2,h_2))\,(V_1 - V_2) \le -g_L(V_1-V_2)^2 &\\ + 4g_K|V_2 - V_K||n_1-n_2||V_1-V_2| &\\ + 3g_{Na}|V_2 - V_{Na}||m_1-m_2||V_1-V_2| &\\ + g_{Na}|V_2 - V_{Na}||h_1-h_2||V_1-V_2|. &\end{aligned} \tag{B.1}$$

for every $m_i, n_i, h_i \in [0,1]$ and $V_i \in \mathbb{R}$, $i = 1,2$.

Proof. Since

$$x^4 - y^4 = (x^2+y^2)(x+y)(x-y),$$

and

$$x^3u - y^3v = u(x^2+xy+y^2)(x-y) + y^3(u-v),$$

we get

$$\begin{aligned}F(V_1,m_1,n_1,h_1) - F(V_2,m_2,n_2,h_2) = &-(g_K n_1^4 + g_{Na} m_1^3 h_1 + g_L)(V_1-V_2)\\ &- g_K(V_2 - V_K)(n_1^2+n_2^2)(n_1+n_2)(n_1-n_2)\\ &- g_{Na}(V_2 - V_{Na})h_1(m_1^2 + m_1m_2 + m_2^2)(m_1-m_2)\\ &- g_{Na}(V_2-V_{Na})m_2^3(h_1-h_2)\end{aligned}$$

and the asserted bound follows. ☐

The following result is the core of the proof of Theorem 2.3:

Proposition B.2. *For each $V^* > 0$, there are constants $J_E^* > 0$ and $\lambda^* > 0$ not depending on N nor on σ such that for each $J_E > J_E^*$ and any solution X of (2.1) satisfying $V_{0,\infty}^{max} \le V^*$, one has*

$$\mathbb{E}\left(|X_t^{(i)} - X_t^{(j)}|^2\right) \le \mathbb{E}\left(|X_0^{(i)} - X_0^{(j)}|^2\right) e^{-\lambda^* t} + \sigma^2 \frac{2C^*_{\zeta,\rho}}{\lambda^*} \quad \forall\, t \ge 0,$$

for all $i,j \in \{1,\dots,N\}$, where

$$C^*_{\zeta,\rho} = \sum_{x=m,n,h,y} \|\rho_x \vee \zeta_x\|_{\infty,V^*} < \infty.$$

Proof. Let us write $\Delta V_t = V_t^{(i)} - V_t^{(j)}$ and $\Delta x_t = x_t^{(i)} - x_t^{(j)}$. Thanks to the bound (B.1), we have

$$\begin{aligned}(\Delta V_t)^2 = &(\Delta V_0)^2 + 2\int_0^t [F(V_s^{(i)}, m_s^{(i)}, n_s^{(i)}, h_s^{(i)}) - F(V_s^{(j)}, m_s^{(j)}, n_s^{(j)}, h_s^{(j)})]\Delta V_s ds\\ &- \int_0^t (2J_E + 2J_{Ch}\bar{y}_s^N)(\Delta V_s)^2 ds\\ \le\ &(\Delta V_0)^2 + \int_0^t 8g_K|V_s^{(j)} - V_K||\Delta n_s||\Delta V_s| + 6g_{Na}|V_s^{(j)} - V_{Na}||\Delta m_s||\Delta V_s| ds\\ &+ \int_0^t 2g_{Na}|V_s^{(j)} - V_{Na}||\Delta h_s||\Delta V_s| ds - \int_0^t (2g_L + 2J_E + 2J_{Ch}\bar{y}_s^N)(\Delta V_s)^2 ds\\ \le\ &(\Delta V_0)^2 + \int_0^t \varepsilon_m (\Delta m_s)^2 + \varepsilon_n (\Delta n_s)^2 + \varepsilon_h (\Delta h_s)^2 ds\\ &- \int_0^t \left(2g_L + 2J_E + 2J_{Ch}\bar{y}_s^N - \frac{9M_{Na}^2}{\varepsilon_m} - \frac{16M_K^2}{\varepsilon_n} - \frac{M_{Na}^2}{\varepsilon_h}\right)(\Delta V_s)^2 ds,\end{aligned}$$

where we have used Young's inequality: $ab \leq \varepsilon_x a^2 + \frac{b^2}{4\varepsilon_x}$ for $x = m, n, h, y$, with $\varepsilon_x > 0$ to be chosen later, and where we have set

$$M_{\text{Na}} = g_{\text{Na}} \max_{v \in [-V^*, V^*]} |v - V_{\text{Na}}|, \quad M_{\text{K}} = g_{\text{K}} \max_{v \in [-V^*, V^*]} |v - V_{\text{K}}|.$$

On the other hand, for the channel types $x = m, n, h, y$, we have

$$\begin{aligned}
\mathbb{E}\left[(\Delta x_t)^2\right] = \mathbb{E}\left[(\Delta x_0)^2\right] &+ 2\int_0^t \mathbb{E}\left[(1 - x_t^{(i)})(\rho_x(V_t^{(i)}) - \rho_x(V_t^{(j)}))\Delta x_s\right] ds \\
&- 2\int_0^t \mathbb{E}\left[x_t^{(i)}(\zeta_x(V_t^{(i)}) - \zeta_x(V_t^{(j)}))\Delta x_s\right] ds \\
&- 2\int_0^t \mathbb{E}\left[\left(\rho_x(V_s^{(j)}) + \zeta_x(V_s^{(j)})\right)(\Delta x_s)^2\right] ds \\
&+ \int_0^t \mathbb{E}\left[\sigma_x^2(V_s^{(j)}, x_s^{(j)}) + \sigma_x^2(V_s^{(i)}, x_s^{(i)})\right] ds.
\end{aligned}$$

By our assumptions, for all $t \geq 0$ we have fror $k = i, j$,

$$\sigma_x^2(V_t^{(k)}, x_t) \leq \sigma^2\left((1 - x_t^{(k)})\rho_x(V_t^{(k)}) + x_t^{(k)}\zeta_x(V_t^{(k)})\right) \leq \sigma^2\|\rho_x \vee \zeta_x\|_{\infty, V^*}.$$

Using Young's inequality in the same way as before yields

$$\begin{aligned}
\mathbb{E}\left[(\Delta x_t)^2\right] \leq \mathbb{E}\left[(\Delta x_0)^2\right] &+ \int_0^t \mathbb{E}\left[\frac{(L^*_{\rho_x} + L^*_{\zeta_x})^2}{\varepsilon_x}(\Delta V_s)^2\right] ds \\
&- (2\eta_x - \varepsilon_x)\int_0^t \mathbb{E}\left[(\Delta x_s)^2\right] ds + 2t\,\sigma^2\|\rho_x \vee \zeta_x\|_{\infty, V^*},
\end{aligned}$$

where L^*_f denotes the Lipschitz constant on $[-V^*, V^*]$ of a locally Lipschitz function f, and where

$$\eta_x^* := \inf_{v \in [-V^*, V^*]} \{\rho_x(v) + \zeta_x(v)\} > 0.$$

Adding up, we get

$$\begin{aligned}
\mathbb{E}\left(|X_t^{(i)} - X_t^{(j)}|^2\right) \leq \mathbb{E}&\left[|X_0^{(i)} - X_0^{(j)}|^2\right] \\
&- \int_0^t \mathbb{E}\Big[\Big(2g_{\text{L}} + 2J_{\text{E}} - \frac{9M_{\text{Na}}^2 + (L^*_{\rho_m} + L^*_{\zeta_m})^2}{\varepsilon_m} - \frac{16M_{\text{K}}^2 + (L^*_{\rho_n} + L^*_{\zeta_n})^2}{\varepsilon_n} \\
&\qquad - \frac{M_{\text{Na}}^2 + (L^*_{\rho_h} + L^*_{\zeta_h})^2}{\varepsilon_h} - \frac{(L^*_{\rho_y} + L^*_{\zeta_y})^2}{\varepsilon_y}\Big)(\Delta V_s)^2\Big] ds \\
&- (2\eta_m - 2\varepsilon_m)\int_0^t \mathbb{E}\left[(\Delta m_s)^2\right] ds - (2\eta_n - 2\varepsilon_n)\int_0^t \mathbb{E}\left[(\Delta n_s)^2\right] ds \\
&- (2\eta_h - 2\varepsilon_h)\int_0^t \mathbb{E}\left[(\Delta h_s)^2\right] ds - (2\eta_y - \varepsilon_y)\int_0^t \mathbb{E}\left[(\Delta y_s)^2\right] ds \\
&+ 2t\,\sigma^2 C^*_{\zeta,\rho}.
\end{aligned}$$

Define now λ^* as the optimal value of the problem

$$\max_{J, \varepsilon_m, \varepsilon_n, \varepsilon_h, \varepsilon_y > 0} \Psi(J, \varepsilon_m, \varepsilon_n, \varepsilon_h, \varepsilon_y),$$

where

$$\begin{aligned}
&\Psi(J, \varepsilon_m, \varepsilon_n, \varepsilon_h, \varepsilon_y) := \\
&\min\Bigg\{2g_{\text{L}} + 2J - \frac{12M_{\text{Na}}^2 + (L^*_{\rho_h} + L^*_{\zeta_h})^2}{\varepsilon_m} - \frac{16M_{\text{K}}^2 + (L^*_{\rho_n} + L^*_{\zeta_n})^2}{\varepsilon_n} - \frac{M_{\text{Na}}^2 + (L^*_{\rho_h} + L^*_{\zeta_h})^2}{\varepsilon_h} - \frac{(L^*_{\rho_y} + L^*_{\zeta_y})^2}{\varepsilon_y}, \\
&\qquad\qquad 2\eta_m^* - 2\varepsilon_m,\, 2\eta_n^* - 2\varepsilon_n, 2\eta_h^* - 2\varepsilon_h, 2\eta_y^* - \varepsilon_y\Bigg\}.
\end{aligned}$$

Notice that λ^* is strictly positive since $\Psi(J,\varepsilon_m,\varepsilon_n,\varepsilon_h,\varepsilon_y)$ can be made so by taking small enough $\varepsilon_x > 0$ for $x = m,n,h,y$ and then large enough $J > 0$. Calling J^*_{E} the smallest $J > 0$ such that $(J,\varepsilon_m,\varepsilon_n,\varepsilon_h,\varepsilon_y) \in \arg\max \Psi$, it follows that for every $J_{\mathrm{E}} > J^*_{\mathrm{E}}$,

$$\mathbb{E}\left(|X_t^{(i)} - X_t^{(j)}|^2\right) \leq \mathbb{E}\left[|X_0^{(i)} - X_0^{(j)}|^2\right] - \lambda^* \int_0^t \mathbb{E}\left(|X_s^{(i)} - X_s^{(j)}|^2\right) + 2t\,\sigma^2 C_{\zeta,\rho}.$$

Applying Lemma A.5, we obtain

$$\sqrt{\mathbb{E}\left(|X_t^{(i)} - X_t^{(j)}|^2\right)} \leq e^{-\frac{\lambda^* t}{2}} \left(\mathbb{E}\left[|X_0^{(i)} - X_0^{(j)}|^2\right] + \frac{(e^{\lambda^* t} - 1)}{\lambda^*} 2\sigma^2 C^*_{\zeta,\rho}\right)^{1/2},$$

and the desired result. □

The next result removes the dependance of the previous one on the bound V^*, at the price of ensuring exponentially fast synchronization only from some time instant $t_0 \geq 0$ on. It will then be easy to deduce part a) of Theorem 2.3.

Theorem B.3. *There are constants $J_E^0 > 0$ and $\lambda^0 > 0$ not depending on $N \geq 1$, on $\sigma \geq 0$ nor on the initial data, and $t_0 \geq 0$ not depending on $N \geq 1$ nor on $\sigma \geq 0$, such that for each $J_E > J_E^0$ the solution X of (2.1) satisfies, for every $t \geq t_0$,*

$$\mathbb{E}\left(|X_t^{(i)} - X_t^{(j)}|^2\right) \leq \mathbb{E}\left(|X_{t_0}^{(i)} - X_{t_0}^{(j)}|^2\right) e^{-\lambda^0 (t-t_0)} + \sigma^2 \frac{2C^0_{\zeta,\rho}}{\lambda^0}, \quad \forall\, i,j \in \{1,\ldots,N\},$$

where

$$C^0_{\zeta,\rho} := \sum_{x=m,n,h,y} \|\rho_x \vee \zeta_x\|_{\infty,\frac{5R_{max}}{g_L}} < \infty\,.$$

Proof. Fix $\epsilon_0 \in (0,1)$, take $t_0 \geq 0$ such that $2V_0^{\max} e^{-g_L t_0} \leq \epsilon_0 \frac{R_{\max}}{g_L}$ and, conditionally on the sigma-field generated by $(X_s : s \leq t_0)$, apply Proposition B.2 to the shifted process $X' := (X_{t+t_0} : t \geq 0)$ with $V^* = V^*_{t_0} \leq (4+\epsilon_0)\frac{R_{\max}}{g_L} \leq 5\frac{R_{\max}}{g_L}$. The proof is then achieved taking expectation in the obtained inequality. □

We can now finish the proof of Theorem 2.3. a). Here and in the sequel we denote by $\bar{S}_t^V$ and $\bar{S}_t^x$ the empirical variance of voltages and x type channels at time t, respectively:

$$\bar{S}_t^V = \frac{1}{N}\sum_{i=1}^N \left(V_t^{(i)} - \bar{V}_t^N\right)^2 \quad \text{and} \quad \bar{S}_t^x = \frac{1}{N}\sum_{i=1}^N \left(x_t^{(i)} - \bar{x}_t^N\right)^2.$$

Proof of Theorem 2.3. a). Applying in the conclusion of Theorem B.3 the elementary identity

$$\frac{1}{N^2}\sum_{i,j=1}^N (\alpha_i - \alpha_j)^2 = \frac{2}{N}\sum_{k=1}^N (\alpha_k - \bar{\alpha}^N)^2 \quad \text{for every } \alpha_1,\ldots,\alpha_N \in \mathbb{R},$$

with $\bar{\alpha}^N = \frac{1}{N}\sum_{i=1}^N \alpha_i$ we get

$$\mathbb{E}\left(\bar{S}_t^V + \sum_{x=mn,n,h,y} \bar{S}_t^x\right) \leq \mathbb{E}\left(\bar{S}_{t_0}^V + \sum_{x=mn,n,h,y} \bar{S}_{t_0}^x\right) e^{-\lambda^0(t-t_0)} + \sigma^2 \frac{C^0_{\zeta,\rho}}{\lambda^0}$$

in the general case. If, additionally, exchangeability of the initial condition is assumed, the path law of system (2.1) is exchangeable, by pathwise uniqueness. The asserted inequality follows. □

Remark B.4.

i) Theorems 2.3 and B.3 show that, for large enough J_E, synchronization of the network (2.1) always occurs, as long as the initial voltage V_0 is bounded, but regardless of its actual values. More precisely, the time $t_0 > 0$ which depends on V_0^{max}, on $\frac{R_{max}}{g_L}$ and on some arbitrary choice of the parameter $\epsilon_0 > 0$, but not on J_E, is one possible time after which we can grant that *the voltage trajectories stay in some fixed interval not depending on V_0. Then, after t_0 and if J_E was chosen large enough, synchronization occurs at least at the exponential rate λ^0 which depends on coefficients of the system (2.1) but no longer on the initial data. In turn, for large enough J_E, Proposition B.2 ensures synchronization from $t_0 = 0$ on but only if V_0^{max} is small enough.*

ii) Notice that the function Ψ in the proof of Proposition B.2 (and hence the constant λ^* therein) increases when its parameter V^* decreases, whereas $C^*_{\zeta,\rho}$ decreases when V^* does. Therefore, letting $\epsilon_0 \to 0$ (or $t_0 \to \infty$) yields the best (by this approach) bounds for the $\limsup$ in Theorem 2.3. Moreover, the largest possible exponential rate $\lambda^0 > 0$ and the smallest possible interaction strength $J^0_E \geq 0$ that can be obtained (but not necessarily attained) in Theorems 2.3 and B.3 by our approach are λ^* and J^*_E corresponding to $V^* = \frac{4R_{max}}{g_L}$. These choices are certainly not optimal in general.

C Synchronized dynamics: proof of Theorem 2.3 *b)*

Our next goal is to prove part b) of Theorem 2.3.

Remark C.1. *Proceeding in a similar way as in the proof of Proposition A.3 one checks that the process (2.6) satisfies $\frac{d}{dt}|\widehat{V}_t|_2^2 + 2g_L|\widehat{V}_t|_2^2 \leq 2R_{max}|\widehat{V}_t|$, which now yields, for any $t \geq t_1$,*

$$|\widehat{V}_t| \leq |\widehat{V}_{t_1}|e^{-g_L(t-t_1)} + \frac{2R_{max}}{g_L}(1 - e^{-g_L(t-t_1)}).$$

Applying on $\bar{V}^N_{t_1} = \widehat{V}_{t_1}$ the first bound in Lemma A.3 we get that $|\widehat{V}_t| \leq V_0^{max}e^{-g_Lt} + \frac{2R_{max}}{g_L}$ for every $t \geq t_1$. Thus, if $t_0 \geq 0$ is chosen as in Theorem B.3, we deduce that

$$\max\left\{\sup_{s\in[t_1,\infty)}|\bar{V}_s|, \sup_{s\in[t_1,\infty)}|\widehat{V}_s|\right\} \leq \frac{V^*_{t_0}}{2} \leq \frac{5R_{max}}{2g_L}. \tag{C.1}$$

We first prove

Proposition C.2. *Let t_0 be as in Theorem B.3 and $\delta > 0$. There are constants $K_{1,\delta}, K_{2,\delta} > 0$ increasingly depending on $\delta > 0$, but not depending on N nor on the initial condition, such that for each $t_1 \geq t_0$,*

$$\mathbb{E}\left(\sup_{t_1\leq t\leq t_1+\delta}|\bar{X}^{N,t_1}_t - \widehat{X}^N_t|^2\right) \leq \left(\left[(V^*_{t_0})^2 + 4\right]e^{-\lambda^0(t_1-t_0)} + \frac{\sigma^2 C^0_{\zeta,\rho}}{\lambda^0}\right)\delta K_{1,\delta} + \delta K_{2,\delta}\frac{\sigma^2}{N}C^0_{\zeta,\rho}. \tag{C.2}$$

Proof. For notational simplicity we write in the proof $\widehat{X}^N_t := \widehat{X}^{N,t_1}_t$. Notice that the average process satisfies the dynamics

$$\bar{V}^N_t = \bar{V}^N_{t_1} + \int_{t_1}^t \frac{1}{N}\sum_{i=1}^N F(V^{(i)}_s, m^{(i)}_s, n^{(i)}_s, h^{(i)}_s) - J_{\text{Ch}}\bar{y}^N_s(\bar{V}^N_s - V_{\text{rev}})ds$$
$$\bar{x}^N_t = \bar{x}^N_{t_1} + \frac{1}{N}\sum_{j=1}^N\int_{t_1}^t \rho_x(V^{(j)}_s)(1 - x^{(j)}_s) - \zeta_x(V^{(j)}_s)x^{(j)}_s + \frac{1}{N}\sum_{i=1}^N\int_{t_1}^t \sigma_x(V^{(j)}_s, x^{(j)}_s)dW^{x,j}_s.$$

Therefore, after some manipulations, we get that

$$\begin{aligned}\left(\bar{V}^N_t - \widehat{V}_t\right)^2 &= \int_{t_1}^t \left[\frac{1}{N}\sum_{i=1}^N F(V^{(i)}_s, m^{(i)}_s, n^{(i)}_s, h^{(i)}_s) - F(\bar{V}^N_s, \bar{m}^N_s, \bar{n}^N_s, \bar{h}^N_s)\right]^2 ds \\ &\quad + 2\int_{t_1}^t \left[F(\bar{V}^N_s, \bar{m}^N_s, \bar{n}^N_s, \bar{h}^N_s) - F(\widehat{V}^N_s, \widehat{m}^N_s, \widehat{n}^N_s, \widehat{h}^N_s)\right](\bar{V}^N_s - \widehat{V}^N_s)ds \\ &\quad + \int_{t_1}^t (1 - 2J_{\text{Ch}}\bar{y}^N_s)(\bar{V}^N_s - \widehat{V}^N_s)^2 + 2J_{\text{Ch}}(\widehat{V}^N_s - V_{\text{rev}})(\bar{y}^N_s - \widehat{y}^N_s)(\bar{V}^N_s - \widehat{V}^N_s)ds \\ &= I_1 + I_2 + I_3.\end{aligned}$$

By Jensen's inequality and the bound (C.1) we have

$$
\begin{aligned}
I_1 &\leq \int_{t_1}^{t} \frac{1}{N}\sum_{i=1}^{N}\left[F(V_s^{(i)}, m_s^{(i)}, n_s^{(i)}, h_s^{(i)}) - F(\bar{V}_s^N, \bar{m}_s^N, \bar{n}_s^N, \bar{h}_s^N)\right]^2 ds\\
&\leq \int_{t_1}^{t} \frac{1}{N}\sum_{i=1}^{N} 4\left[\left(g_{\mathrm{K}}(n_s^{(i)})^4 + g_{\mathrm{Na}}(m_s^{(i)})^3 h_s^{(i)} + g_{\mathrm{L}}\right)(V_s^{(i)} - \bar{V}_s^N)\right]^2 ds\\
&\quad + \int_{t_1}^{t} \frac{1}{N}\sum_{i=1}^{N} 4\left[g_{\mathrm{Na}}(\bar{V}_s^N - V_{\mathrm{Na}})h_s^{(i)}\left((m_s^{(i)})^2 + m_s^{(i)}\bar{m}_s^N + (\bar{m}_s^N)^2\right)(m_s^{(i)} - \bar{m}_s^N)\right]^2 ds\\
&\quad + \int_{t_1}^{t} \frac{1}{N}\sum_{i=1}^{N} 4\left[g_{\mathrm{Na}}(\bar{V}_s^N - V_{\mathrm{K}})\left((n_s^{(i)})^2 + (\bar{n}_s^N)^2\right)\left(n_s^{(i)} + \bar{n}_s^N\right)(n_s^{(i)} - \bar{n}_s^N)\right]^2 ds\\
&\quad + \int_{t_1}^{t} \frac{1}{N}\sum_{i=1}^{N} 4\left[g_{\mathrm{Na}}(\bar{V}_s^N - V_{\mathrm{Na}})(\bar{m}_s^N)^3(h_s^{(i)} - \bar{h}_s^N)\right]^2 ds\\
&\leq K_V^1 \int_{t_1}^{t} \bar{S}_s^V + \bar{S}_s^m + \bar{S}_s^n + \bar{S}_s^h ds,
\end{aligned}
$$

with K_V^1 explicitly depending on $\sup_{v\in[-\frac{5R_{\max}}{2g_L}, \frac{5R_{\max}}{2g_L}]} \max\{|v - V_{\mathrm{Na}}|, |v - V_{\mathrm{K}}|\}$, g_{K} and g_{Na}. Meanwhile, using (B.1) we get

$$
\begin{aligned}
I_2 &\leq \int_{t_1}^{t} -2g_{\mathrm{L}}(\bar{V}_s^N - \widehat{V}_s^N)^2 + 4g_{\mathrm{K}}|\widehat{V}_s^N - V_{\mathrm{K}}|\left((\bar{V}_s^N - \widehat{V}_s^N)^2 + (\bar{n}_s^N - \widehat{n}_s^N)^2\right)\\
&\quad + 3g_{\mathrm{Na}}|\widehat{V}_s^N - V_{\mathrm{Na}}|\left((\bar{V}_s^N - \widehat{V}_s^N)^2 + (\bar{m}_s^N - \widehat{m}_s^N)^2\right)\\
&\quad + g_{\mathrm{Na}}|\widehat{V}_s^N - V_{\mathrm{Na}}|\left((\bar{V}_s^N - \widehat{V}_s^N)^2 + (\bar{h}_s^N - \widehat{h}_s^N)^2\right) ds\\
&\leq K_V^2 \int_{t_1}^{t} (\bar{V}_s^N - \widehat{V}_s^N)^2 + (\bar{n}_s^N - \widehat{n}_s^N)^2 + (\bar{m}_s^N - \widehat{m}_s^N)^2 + (\bar{h}_s^N - \widehat{h}_s^N)^2 ds,
\end{aligned}
$$

with K_V^2 also depending on those quantities and on g_L. By similar arguments, we get

$$
I_3 \leq K_V^3 \int_{t_1}^{t} (\bar{V}_s^N - \widehat{V}_s^N)^2 + (\bar{y}_s^N - \widehat{y}_s^N)^2 ds
$$

for some K_V^3 depending on J_{Ch} and on $\sup_{v\in[-\frac{5R_{\max}}{2g_L}, \frac{5R_{\max}}{2g_L}]} |v - V_{\mathrm{rev}}|$. We thus get:

$$
\left(\bar{V}_t^N - \widehat{V}_t\right)^2 \leq K_V^1 \int_{t_1}^{t} \left[\bar{S}_s^V + \bar{S}_s^m + \bar{S}_s^n + \bar{S}_s^h\right] ds + \tilde{K}_V \int_{t_1}^{t} |\bar{X}_s^N - \widehat{X}_s^N|^2 ds
$$

for some explicit $\tilde{K}_V$ a.s., from where

$$
\mathbb{E}\left[\sup_{t_1\leq s\leq t}\left(\bar{V}_s^N - \widehat{V}_s\right)^2\right] \leq K_V^1 \int_{t_1}^{t} \mathbb{E}\left[\bar{S}_s^V + \bar{S}_s^m + \bar{S}_s^n + \bar{S}_s^h\right] ds + \tilde{K}_V \int_{t_1}^{t} \mathbb{E}\left[\sup_{t_1\leq u\leq s} |\bar{X}_u^N - \widehat{X}_u^N|^2\right] ds. \tag{C.3}
$$

On the other hand, for x type channels we get

$$
\begin{aligned}
\bar{x}_t^N - \widehat{x}_t^N &= \frac{1}{N}\sum_{j=1}^{N}\int_{t_1}^{t} \rho_x(V_s^{(j)})(1 - x_s^{(j)}) - \zeta_x(V_s^{(j)})x_s^{(j)} - \rho_x(\bar{V}_s^N)(1 - \bar{x}_s^N) + \zeta_x(\bar{V}_s^N)\bar{x}_s^N ds\\
&\quad + \int_{t_1}^{t} \rho_x(\bar{V}_s^N)(1 - \bar{x}_s^N) - \zeta_x(\bar{V}_s^N)\bar{x}_s^N - \rho_x(\widehat{V}_s^N)(1 - \widehat{x}_s^N) + \zeta_x(\widehat{V}_s^N)\widehat{x}_s^N ds\\
&\quad + \frac{1}{N}\sum_{j=1}^{N}\int_{t_1}^{t} \sigma_x(V_s^{(j)}, x_s^{(j)}) dW_s^{x,j}.
\end{aligned}
$$

For $t \in (t_1, t_1+\delta)$ we deduce:

$$
\begin{aligned}
(\bar{x}_t^N - \widehat{x}_t^N)^2 \leq 3\delta \int_{t_1}^t \frac{1}{N}\sum_{j=1}^N \Big(\rho_x(V_s^{(j)})(1-x_s^{(j)}) - \zeta_x(V_s^{(j)})x_s^{(j)} - \rho_x(\bar{V}_s^N)(1-\bar{x}_s^N) + \zeta_x(\bar{V}_s^N)\bar{x}_s^N\Big)^2\, ds \\
+ 3\delta \int_0^t \Big(\rho_x(\bar{V}_s^N)(1-\bar{x}_s^N) - \zeta_x(\bar{V}_s^N)\bar{x}_s^N - \rho_x(\widehat{V}_s^N)(1-\widehat{x}_s^N) + \zeta_x(\widehat{V}_s^N)\widehat{x}_s^N\Big)^2\, ds \\
+ 3\left(\frac{1}{N}\sum_{i=1}^N \int_{t_1}^t \sigma_x(V_s^{(j)}, x_s^{(j)})dW_s^{x,j}\right)^2.
\end{aligned}
$$

The previous yields,

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{t_1\leq s\leq t} (\bar{x}_s^N - \widehat{x}_s^N)^2\Big] \\
&\leq 3\delta \int_{t_1}^t \mathbb{E}\left[\frac{1}{N}\sum_{j=1}^N \Big(\rho_x(V_s^{(j)})(1-x_s^{(j)}) - \zeta_x(V_s^{(j)})x_s^{(j)} - \rho_x(\bar{V}_s^N)(1-\bar{x}_s^N) + \zeta_x(\bar{V}_s^N)\bar{x}_s^N\Big)^2\right] ds \\
&\quad + 3\delta \int_0^t \mathbb{E}\left[\Big(\rho_x(\bar{V}_s^N)(1-\bar{x}_s^N) - \zeta_x(\bar{V}_s^N)\bar{x}_s^N - \rho_x(\widehat{V}_s^N)(1-\widehat{x}_s^N) + \zeta_x(\widehat{V}_s^N)\widehat{x}_s^N\Big)^2\right] ds \\
&\quad + 3\mathbb{E}\left[\sup_{t_1\leq s\leq t}\left(\frac{1}{N}\sum_{i=1}^N \int_{t_1}^s \sigma_x(V_u^{(j)}, x_u^{(j)})dW_u^{x,j}\right)^2\right] \\
&= I_1 + I_2 + I_3.
\end{aligned}
$$

Denoting by $L_{f,R}$ a Lipschitz constant of a function f on $[-R,R]$ and using standard arguments, we get that

$$
I_1 \leq K_x\delta(L^2_{\rho_x, \frac{5R_{\max}}{2g_L}} + L^2_{\rho_x+\zeta_x, \frac{5R_{\max}}{2g_L}}) \int_{t_1}^t \mathbb{E}\left[\bar{S}_s^V + \bar{S}_s^x\right] ds
$$

and that

$$
I_2 \leq K_x\delta(L^2_{\rho_x, \frac{5R_{\max}}{2g_L}} + L^2_{\rho_x+\zeta_x, \frac{5R_{\max}}{2g_L}}) \int_{t_1}^t \mathbb{E}\left[(\bar{V}_s^N - \widehat{V}_s^N)^2 + (\bar{x}_s^N - \widehat{x}_s^N)^2\right] ds
$$

for all $t \in (t_1, t_1+\delta)$. By Doob's inequality, we moreover obtain

$$
I_3 \leq 3\cdot 4\mathbb{E}\left[\frac{1}{N^2}\sum_{i=1}^N \int_{t_1}^t \sigma_x^2(V_s^{(j)}, x_s^{(j)})ds\right] \leq \frac{12\sigma^2\delta}{N}\|\rho_x \vee \zeta_x\|_{\infty, \frac{5R_{\max}}{2g_L}}.
$$

Summarizing, for the x-type channel we have shown that for all $t \in (t_1, t_1+\delta)$,

$$
\mathbb{E}\left[\sup_{t_1\leq s\leq t} (\bar{x}_s^N - \widehat{x}_s^N)^2\right] \leq \delta K_x \int_{t_1}^t \mathbb{E}\left[\bar{S}_s^V + \bar{S}_s^x + \sup_{t_1\leq u\leq s}|\bar{X}_u^N - \widehat{X}_u^N|^2\right] ds + \frac{12\sigma^2\delta}{N}\|\rho_x \vee \zeta_x\|_{\infty, \frac{5R_{\max}}{2g_L}} \tag{C.4}
$$

for some constants $K_x > 0$. Putting together (C.3) and (C.4) we get for all $t \in (t_1, t_1+\delta)$ and some constants $K_1, K_2 > 0$,

$$
\begin{aligned}
\mathbb{E}\left[\sup_{t_1\leq s\leq t}|\bar{X}_s^N - \widehat{X}_s^N|^2\right] \leq (1+\delta)K_1 \int_{t_1}^{t_1+\delta} \mathbb{E}\left(\bar{S}_s^V + \sum_{x=m,n,h,y}\bar{S}_s^x\right) ds + \frac{12\sigma^2\delta}{N}C^0_{\zeta,\rho} \\
+ (1+\delta)K_2 \int_{t_1}^t \mathbb{E}\left[\sup_{t_1\leq u\leq s}|\bar{X}_u^N - \widehat{X}_u^N|^2\right] ds,
\end{aligned}
$$

from where, using Gronwall's inequality, we deduce:

$$
\mathbb{E}\left(\sup_{t_1\leq t\leq t_1+\delta}|\bar{X}_t^N - \widehat{X}_t^N|^2\right) \leq e^{K_2(1+\delta)}\left(K_1(1+\delta)\int_{t_1}^{t_1+\delta}\mathbb{E}\left(\bar{S}_s^V + \sum_{x=m,n,h,y}\bar{S}_s^x\right) ds + \frac{12\sigma^2\delta}{N}C^0_{\zeta,\rho}\right).
$$

We can now use Theorem 2.3 to bound the integral on the r.h.s. With $K_{1,\delta} = e^{K_2(1+\delta)}K_1(1+\delta)$ and $K_{1,\delta} = 12e^{K_2(1+\delta)}$ we get, for all $t_1 \geq t_0$, that

$$\begin{aligned}\mathbb{E}\left(\sup_{t_1\le t\le t_1+\delta}|\bar{X}^N_t - \widehat{X}^N_t|^2\right) &\le \mathbb{E}\left(\bar{S}^V_{t_0} + \sum_{x=m,n,h,y}\bar{S}^x_{t_0}\right)\frac{1}{\lambda^0}(1-e^{-\lambda^0\delta})e^{-\lambda^0(t_1-t_0)}K_{1,\delta} \\ &\quad + \frac{\sigma^2 C^0_{\zeta,\rho}}{\lambda^0}\delta K_{1,\delta} + \delta K_{2,\delta}\frac{\sigma^2}{N}C^0_{\zeta,\rho} \\ &\le \left(\left[(V^*_{t_0})^2+4\right]e^{-\lambda^0(t_1-t_0)} + \frac{\sigma^2 C^0_{\zeta,\rho}}{\lambda^0}\right)\delta K_{1,\delta} + \delta K_{2,\delta}\frac{\sigma^2}{N}C^0_{\zeta,\rho}\end{aligned}$$

since $\bar{S}^V_{t_0} \le (V^*_{t_0})^2$. □

Proof of Theorem 2.3. b). Notice on hand that, for each $t \geq t_1$, we always have the bounds

$$|X^{(i)}_t - \widehat{X}^{t_1,N}_t|^2 \le 2\bar{S}^V_t + 2|\widehat{V}_t - \bar{V}^N_t|^2 + 4 \le 4(V^*_{t_0})^2 + 4 \le K_0 := 4\left(\frac{5R_{\max}}{g_L}\right)^2 + 4,$$

thanks to (C.1) and that $V^*_{t_0} \le \frac{5R_{\max}}{g_L}$. On the other hand, combining Proposition C.2 with Theorem 2.3. a) we get for every $t \in [t_1, t_1+\delta]$ that

$$\mathbb{E}\left(|X^{(i)}_t - \widehat{X}^{t_1,N}_t|^2\right) \le 2\left[\left(K'_0 e^{-\lambda_0(t-t_0)} + \sigma^2\frac{C^0_{\zeta,\rho}}{\lambda^0}\right)(1+\delta K_{1,\delta}) + \delta K_{2,\delta}\frac{\sigma^2}{N}C^0_{\zeta,\rho}\right]$$

with $K'_0 = \left(\frac{5R_{\max}}{g_L}\right)^2 + 4$. The statement follows. □

D Propagation of Chaos and synchronization for the McKean-Vlasov limit: proofs of Theorem 2.5 and Corollary 2.7

We first address the asymptotic behavior of the flow of empirical measures (2.9) when $N \to \infty$ and the proof of Theorem 2.5. In particular, we will prove the propagation of chaos property for system (2.1). Following the classic pathwise approach developed in [49] and [39], we first establish:

Theorem D.1. *Under the assumptions of Theorem 2.5, we have:*

a) Let $W^x, x = m,n,h,y$ be independent standard Brownian motions and $(V_0, m_0, n_0, h_0, y_0)$ an independent random vector with law μ_0. There is existence and uniqueness, pathwise and in law, of a solution $\widetilde{X} = (\widetilde{V}_t, \widetilde{m}_t, \widetilde{n}_t, \widetilde{h}_t, \widetilde{y}_t, t \geq 0)$ to the nonlinear stochastic differential equation (in the sense of McKean) with values in $\mathbb{R}\times[0,1]^4$:

$$\begin{aligned}\widetilde{V}_t &= V_0 + \int_0^t F(\widetilde{V}_s, \widetilde{m}_s, \widetilde{n}_s, \widetilde{h}_s)ds - \int_0^t J_E(\widetilde{V}_s - \mathbb{E}[\widetilde{V}_s])ds - \int_0^t J_{Ch}\mathbb{E}[\widetilde{y}_s](\widetilde{V}_s - V_{rev})ds, \\ \widetilde{x}_t &= x_0 + \int_0^t \rho_x(\widetilde{V}_s)(1-\widetilde{x}_s) - \zeta_x(\widetilde{V}_s)\widetilde{x}_s ds + \int_0^t \sigma_x(\widetilde{V}_s, \widetilde{x}_s)dW^x_s, \quad x = m,n,h,y\end{aligned} \tag{D.1}$$

such that for all $t \geq 0$, $|\widetilde{V}_t| \le 4R_{max}/g_L + 2V_0^{max}e^{-g_L t}$ almost surely.

b) $(\mu_t := law(\tilde{X}_t) : t \geq 0)$ is a weak solution globally defined in $C((0,+\infty]; \mathcal{P}_2(\mathbb{R}\times[0,1]^4))$ of the McKean-Vlasov equation (2.10).

c) For each $T > 0$, let $\widetilde{X}^{(i)} = \Big((\widetilde{V}^{(i)}_t, \widetilde{m}^{(i)}_t, \widetilde{n}^{(i)}_t, \widetilde{h}^{(i)}_t, \widetilde{y}^{(i)}_t) : t \in [0,T]\Big)$, $i = 1,\dots,N$ be independent copies of the nonlinear process (D.1) each of them driven by the same Brownian motions $(W^{x,i},\ x = m,n,h,y)$ and with same initial conditions $X^{(i)}_0 = \widetilde{X}^{(i)}_0$ as the N-particle system (2.1). Then, there is a constant $C(T) > 0$ such that for every $N \geq 1$ and $i \in \{1,\dots,N\}$,

$$\mathbb{E}\left[\sup_{0\le t\le T}|X^{(i)}_t - \widetilde{X}^{(i)}_t|^2\right] \le \frac{C(T)}{N}.$$

Proof. The statements a), b) and c) would be standard if the coefficients in each of the N components of (2.1) were replaced by globally Lipschitz functions of $X_s^{(i)}$ and $X_s^{(j)}$, see Theorems 2.2 and 2.3 in [39]. In particular, with functions p_M^j and F_M defined for fixed $M > 0$ as in Lemma A.1, for any $T > 0$ there is existence and uniqueness, pathwise and in law, of a solution to the nonlinear stochastic differential equation on $[0, T]$:

$$\begin{aligned} \widetilde{V}_t^M &= V_0 + \int_0^t F_M(\widetilde{V}_s^M, \widetilde{m}_s^M, \widetilde{n}_s^M, \widetilde{h}_s^M) ds - \int_0^t J_{\mathrm{E}}(\widetilde{V}_s^M - \mathbb{E}[\widetilde{V}_s^M]) ds \\ &\quad - \int_0^t J_{\mathrm{Ch}} \mathbb{E}[p_M^1(\widetilde{y}_s)](p_M^1(\widetilde{V}_s^M) - V_{\mathrm{rev}}) ds, \\ \widetilde{x}_t^M &= x_0 + \int_0^t \rho_x(p_M^1(\widetilde{V}_s^M)(1 - p_M^1(\widetilde{x}_s^M)) - \zeta_x(p_M^2(\widetilde{V}_s^M)) p_M^1(\widetilde{x}_s^M) ds \\ &\quad + \int_0^t \sigma_x(p_M^1(\widetilde{V}_s^M), \widetilde{x}_s^M) dW_s^x, \quad x = m, n, h, y. \end{aligned} \tag{D.2}$$

Moreover, letting $\widetilde{X}^{(i,M)} = \Big((\widetilde{V}_t^{(i,M)}, \widetilde{m}_t^{(i,M)}, \widetilde{n}_t^{(i,M)}, \widetilde{h}_t^{(i,M)}, \widetilde{y}_t^{(i,M)}) : t \in [0, T] \Big)$, $i = 1, \dots, N$ be independent copies of the nonlinear process (D.2) driven by the same Brownian motions ($W^{x,i}$, $x = m, n, h, y$) and with same initial conditions $X_0^{(i)} = \widetilde{X}_0^{(i)}$ as the system $(X^{(1,M)}, \dots, X^{(N,M)})$ defined in (A.1), we obtain that

$$\mathbb{E}\left[\sup_{0 \le t \le T} |X_t^{(i,M)} - \widetilde{X}_t^{(i,M)}|^2 \right] \le \frac{C_M(T)}{N}$$

for every $N \ge 1$ and $i \in \{1, \dots, N\}$, and some constant $C_M(T) > 0$.

We notice now that, by Proposition A.3, for $M > 0$ large enough the system $(X^{(1)}, \dots, X^{(N)})$ is also a solution to the system of equations (A.1). Pathwise uniqueness of the latter yields for all such $M > 0$ that $(X^{(1)}, \dots, X^{(N)}) = (X^{(1,M)}, \dots, X^{(N,M)})$ on $[0, T]$, from where

$$\mathbb{E}\left[\sup_{0 \le t \le T} |X_t^{(i)} - \widetilde{X}_t^{(i,M)}|^2 \right] \le \frac{C_M(T)}{N} \tag{D.3}$$

for every $N \ge 1$ and $i \in \{1, \dots, N\}$. Furthermore, for any $M' > 0$

$$\mathbb{P}\left(\sup_{0 \le t \le T} \widetilde{x}_t^{(i,M)} \ge M' + \varepsilon \right) \le \mathbb{P}\left(\sup_{0 \le t \le T} x_t^{(i,M)} \ge M' \right) + \frac{2C_M(T)}{N\varepsilon^2}$$

Taking $M' = 1$, letting $N \to \infty$ and then $\varepsilon \to 0$ we deduce that $\widetilde{x}_t^{(i,M)} \le 1$ a.s. for every $t \in [0.T]$ and $i \in \mathbb{N}$. In a similar way, $\widetilde{x}_t^{(i,M)} \ge 0$ and $|\widetilde{V}_t^{(i,M)}| \le V_{t,\infty}^{\max}$ hold a.s. for every $t \in [0, T]$ and $i \in \mathbb{N}$. This implies that for $M > 0$ large enough but fixed, a solution to (D.2) also solves (D.1), and proves the existence part in a).

We show now that any solution have uniform in time bounded compact support, from which uniqueness in part a) will immediately follow. We shall first consider a solution $(U_t, q_t^m, q_t^n, q_t^h, q_t^y)$ of (D.1) with explosion time ξ, and we will show that it coincides with $(\widetilde{V}_t^M, \widetilde{m}_t^M, \widetilde{n}_t^M, \widetilde{h}_t^M, \widetilde{y}_t^M, t \ge 0)$ for a M big enough. For $M > 1$, we define $\tau_M = \inf\{t \ge 0 : \max\{|U_t|, |q_t^m|, |q_t^n|, |q_t^h|, |q_t^y|\} \ge M\}$. Then we observe that the coefficients of (D.1) applied to $(U_t, q_t^m, q_t^n, q_t^h, q_t^y, 0 \le t \le \tau_M)$ coincide with the truncated coefficients of (D.2) and thanks to the uniqueness property for (D.2) we conclude that almost surely

$$(U, q^m, q^n, q^h, q^y)_{t \wedge \tau_M} = (\widetilde{V}^M, \widetilde{m}^M, \widetilde{n}^M, \widetilde{h}^M, \widetilde{y}^M)_{t \wedge \tau_M}.$$

In particular, we observe that $q_{t \wedge \tau_M}^x \in [0, 1]$ for $x = m, n, h, y$, and that $\tau_M = \inf\{t \ge 0 : |U_t| \ge M\}$ for $M > 1$. Moreover the second order moment $\mathbb{E}[U_{t \wedge \tau_M}^2]$ is uniformly bounded in M, since

$$\begin{aligned} U_{t \wedge \tau_M}^2 &= V_0^2 + 2 \int_0^{t \wedge \tau_M} U_s F(U_s, q_s^m, q_s^n, q_s^h) ds - 2 \int_0^{t \wedge \tau_M} J_{\mathrm{E}} U_s (U_s - \mathbb{E}[U_s]) ds \\ &\quad - 2 \int_0^{t \wedge \tau_M} J_{\mathrm{Ch}} \mathbb{E}[q_s^y] U_s (U_s - V_{\mathrm{rev}}) ds, \end{aligned}$$

from where, it is easy to show that

$$\mathbb{E}(U^2_{t\wedge\tau_M}) \leq C_1 + C_2\int_0^t \mathbb{E}(U^2_{s\wedge\tau_M}),$$

and therefore, thanks to Gronwall's inequality

$$\mathbb{E}(U^2_{t\wedge\tau_M}) \leq C_1 e^{C_2 t}.$$

On the other hand $\mathbb{E}(U^2_{t\wedge\tau_M}) = \mathbb{E}(U_t^2 \mathbb{1}_{\tau_M>t}) + M^2\mathbb{P}(\tau_M \leq t)$ and then we can conclude for all $t \geq 0$ and all $M \geq 1$

$$\mathbb{P}(\tau_M \leq t) \leq \frac{C_1 e^{C_2 t}}{M^2}.$$

Since $\tau_M \nearrow \xi$, we conclude that for all t $\mathbb{P}(\xi \leq t) = 0$, from where it follows that ξ is almost surely infinite.

Now, since $(U_t, q_t^m, q_t^n, q_t^h, q_t^y)$ has no explosion, we apply Proposition 3.3 in [10] to get that almost surely $q_t^x \in [0,1]$ for any $t > 0$. Using this, we derive a more precise bound for the second order moment:

$$\mathbb{E}(U_t^2) \leq \mathbb{E}(V_0^2) + 2\int_0^t \sqrt{\mathbb{E}(R_s^2)}\sqrt{\mathbb{E}(U_s^2)} - g_{\mathrm{L}}\mathbb{E}(U_s^2)ds,$$

where as in the proof of Proposition A.3,

$$R_s \leq R_{\max} := \max_{a,b,c\in[0,1]} |I + g_{\mathrm{L}}V_{\mathrm{L}} + g_{\mathrm{K}}V_{\mathrm{K}}a + g_{\mathrm{Na}}V_{\mathrm{Na}}b + J_{\mathrm{Ch}}V_{\mathrm{rev}}c|.$$

Applying one more time Lemma A.5 we conclude

$$\sqrt{\mathbb{E}(U_t^2)} \leq \sqrt{\mathbb{E}(V_0^2)}e^{-g_{\mathrm{L}}t} + \frac{2R_{\max}}{g_{\mathrm{L}}}(1 - e^{-g_{\mathrm{L}}t}).$$

Thus, the second moment of any solution of (D.1) is uniformly bounded in time. Moreover, since the initial condition V_0 is bounded, proceeding exactly as in the proof of Proposition A.3 we obtain that

$$|U_t| \leq \frac{4R_{\max}}{g_{\mathrm{L}}} + 2V_0^{\max}e^{-g_{\mathrm{L}}t},$$

with the same bound $V_0^{\max}$ for V_0. In conclusion, solutions of (D.1) are non explosive, even more they are uniformly bounded in time. Choosing $M > 4R_{\max}/g_{\mathrm{L}} + 2V_0^{\max}$, we get $\tau_M = \infty$ almost surely, and for any $t \geq 0$,

$$(U, q^m, q^n, q^h, q^y)_t = (\widetilde{V}^M, \widetilde{m}^M, \widetilde{n}^M, \widetilde{h}^M, \widetilde{y}^M)_t.$$

Hence equation (D.1) has a unique solution.

Part b) derives from a direct and easy application of the Ito's formula to compute

$$\mathbb{E}[\psi(\widetilde{X}_t)] = \int_{\mathbb{R}\times[0,1]^4} \psi(x)\mu_t(dx)$$

for a C_c^∞ test function ψ, thanks to the fact that the Lebesgue integrals on the right hand side of the Itô formula will be all bounded, since the supports of the laws $(\mu_t : t \geq 0)$ are contained in some compact set, and by continuity of coefficients.

Part c) is immediate taking large enough M in (D.3).

□

We are now in position to prove

Proof of of Theorem 2.5. a) We write $\mathcal{C}_T := C([0,T], \mathbb{R}\times[0,1]^4)$. Part c) of Theorem D.1 implies that for each $T > 0$ and $k \geq 1$ the convergence Law$(X^{(1)}, \dots X^{(k)}) \to \mu^{\otimes k}$ with $\mu = $ Law$(\tilde{X}^{(1)})$ holds on the space $\mathcal{C}_T^k$ as $N \to \infty$. By Proposition 2.2. in [49] or Proposition 4.2. in [39], this implies that the empirical measure

$$\mu^N := \frac{1}{N}\sum_{i=1}^N \delta_{X^{(i)}} \in \mathcal{P}(\mathcal{C}_T),$$

with $\mathcal{P}(\mathcal{C}_T)$ denoting the space of probability measures on $\mathcal{C}_T$ endowed with the weak topology, converges in law to the (deterministic) probability measure μ. The first assertion of the theorem follows then from the fact that the mapping associating with $\nu \in \mathcal{P}(\mathcal{C}_T)$ its flow $(\nu_t : t \in [0,T]) \in C([0,T];\mathcal{P}(\mathbb{R}\times[0,1]^4))$ of one-dimensional time-marginals laws is continuous, together with part b) of Theorem D.1 (notice that $C([0,T];\mathcal{P}(\mathbb{R}\times[0,1]^4))$ can be replaced by $C([0,T];\mathcal{P}_2(\mathbb{R}\times[0,1]^4))$ since all the random measures involved have a common compact support).

b) We observe first that for each $t \geq 0$ one has

$$\mathbb{E}\left(\mathcal{W}_2^2(\mu_t^N,\mu_t)\right) \leq 2\mathbb{E}\left(\mathcal{W}_2^2(\mu_t^N,\tilde{\mu}_t^N)\right) + 2\mathbb{E}\left(\mathcal{W}_2^2(\tilde{\mu}_t^N,\mu_t)\right),$$

where $\tilde{\mu}_t^N$ is the empirical measure of any random i.i.d. sample of the law μ_t constructed in the same probability as μ_t^N. Taking $\tilde{\mu}_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{\tilde{X}_t^{(i)}}$ with $\tilde{X}_t^{(i)}$, $i=1,\dots,N$ the processes defined in part c) of Theorem D.1 we get for every $t\in[0,T]$ that

$$\mathbb{E}\left(\mathcal{W}_2^2(\mu_t^N,\mu_t)\right) \leq 2\frac{C(T)}{N} + 2\mathbb{E}\left(\mathcal{W}_2^2(\tilde{\mu}_t^N,\mu_t)\right).$$

On the other hand, we have $\sup_{t\in[0,T]}(\int |z|^q \mu_t(dz))^{1/q} < \infty$ for each $q \geq 1$, using for instance the bound obtained at the end of the proof of Theorem D.1. We can therefore apply Theorem 1 in [20] with $p=2$, $d=5$ and a sufficiently large $q>2$, to get that $\mathbb{E}\left(\mathcal{W}_2^2(\tilde{\mu}_t^N,\mu_t)\right) \leq CN^{-2/5}$. The second assertion thus follows.

c) In order to prove uniqueness for the McKean-Vlasov equation (2.10), we adapt to our setting a generic argument going back at least to Gärtner [23]. Assume for a while that for each compactly supported $\nu_0 \in \mathcal{P}(\mathbb{R}\times[0,1]^4))$ and $(\nu_t^* : t\in[0,T]) \in C([0,T],\mathcal{P}_2(\mathbb{R}\times[0,1]^4))$ the linear Fokker-Planck equation

$$\partial_t \nu_t = \partial_v\left(\Phi(\langle(\nu_t^*)^V\rangle, \langle(\nu_t^*)^y\rangle,\cdot,\cdot)\nu_t\right) + \sum_{x=m,n,h,y} \frac{1}{2}\sigma^2\,\partial^2_{u_x u_x}(a_x\nu_t) - \partial_{u_x}(b_x\nu_t) \tag{D.4}$$

has at most one weak solution with supports bounded uniformly in $t\in[0,T]$. By similar arguments as in Lemma A.1, strong well-posedness holds for the stochastic differential equation:

$$\begin{aligned} V_t^* &= V_0^* + \int_0^t F(V_s^*, m_s^*, n_s^*, h_s^*)ds - \int_0^t J_{\mathrm{E}}(V_s^* - \langle(\nu_s^*)^V\rangle)ds - \int_0^t J_{\mathrm{Ch}}\langle(\nu_s^*)^y\rangle(V_s^* - V_{\mathrm{rev}})ds, \\ x_t^* &= x_0^* + \int_0^t \rho_x(V_s^*)(1-x_s^*) - \zeta_x(V_s^*)x_s^* ds + \int_0^t \sigma_x(V_s^*, x_s^*)dW_s^x, \quad x=m,n,h,y, \end{aligned} \tag{D.5}$$

with $(V_0^*, m_0^*, n_0^*, h_0^*, y_0^*)$ independent of the Brownian motions W^x and with law ν_0. Moreover, one can check that $x_t^* \in [0,1]$ a.s. for all $t\in[0,T]$ and that the process $(V_t^* : t\in[0,T])$ is bounded. It follows using Itô's formula that a unique weak solution to equation (D.4) with uniformly bounded supports does exist, and is given by $\nu_t = \text{law}(V_t^*, m_t^*, n_t^*, h_t^*, y_t^*)$ for all $t\in[0,T]$. Now, any solution $(\mu_t : t\in[0,T])$ in $C([0,T],\mathcal{P}(\mathbb{R}\times[0,1]^4))$ of (2.10) with uniformly bounded supports also solves the linear equation (D.4) with $(\nu_t^* : t\in[0,T]) = (\mu_t : t\in[0,T])$. This yields, for all $t\in[0,T]$, that $\mu_t = \text{law}(V_t^*, m_t^*, n_t^*, h_t^*, y_t^*)$, for the process defined as in (D.5), with $\nu_s^* = \mu_s$ for all $s\in[0,T]$. In other words, this process solves the nonlinear stochastic differential equation (D.1). From Theorem D.1 we conclude that $(\mu_t : t\in[0,T]) = (\text{law}(\tilde{X}_t) : t\in[0,T])$, that is, there is uniqueness of solutions in $C([0,T],\mathcal{P}(\mathbb{R}\times[0,1]^4))$ of (2.10) having uniformly bounded support.

Hence, in order to conclude the proof of Theorem 2.5 it is enough to show that, given functions $\alpha,\beta \in C([0,T],\mathbb{R})$ and $\nu_0 \in \mathcal{P}_2(\mathbb{R}\times[0,1]^4)$ there is at most one solution $(\nu_t : t\in[0,T]) \in C([0,T],\mathcal{P}(\mathbb{R}\times[0,1]^4))$ with support bounded uniformly in $[0,T]$, to the distribution formulation of equation (D.4)

$$\begin{aligned} \int \psi(t,v,u)\nu_t(dv,du) = \int \psi(0,v,u)\nu_0(dv,du) - \int_0^t \int \Big[& \Phi(\alpha_s,\beta_s,v,u)\partial_v\psi(s,v,u) \\ & + \Big(\partial_s + \sum_{x=m,n,h,y}\frac{1}{2}\sigma^2\, a_x\partial^2_{u_x u_x} + b_x\partial_{u_x}\Big)\psi(s,v,u)\Big]\nu_s(dv,du)\,ds \end{aligned} \tag{D.6}$$

for all $t \in [0,T]$ and for an extended class of test function $\psi \in C_b^{1,1,2}([0,T] \times \mathbb{R} \times [0,1]^4)$. Let ρ'_x and ζ'_x denote compactly supported functions coinciding with ρ_x and ζ_x on some compact set $\mathcal{K} \subset \mathbb{R}$ containing the supports of the measures ν_t^V for $t \in [0,T]$, and define σ'_x, a'_x and b'_x in terms of them in a similar way as σ_x, a_x and b_x were defined in terms of ρ_x and ζ_x. For a given $t > 0$, consider the following Cauchy problem in $\mathbb{R}^5$: for all $(s,v,u) \in [0,t) \times \mathbb{R} \times \mathbb{R}^4$,

$$\begin{aligned} \big(\partial_s - \Phi(\alpha_s, \beta_s, v, u)\partial_v + \sum_{x=m,n,h,y} \frac{1}{2}\sigma^2\, a'_x \partial^2_{u_x u_x} + b'_x \partial_{u_x}\big) f_t(s,v,u) &= 0, \\ f_t(t,v,u) &= \psi(v,u). \end{aligned} \tag{D.7}$$

By the Feynman-Kac formula (see e.g. Karatzas and Shreve [33]), if a solution $f_t \in C_b([0,t] \times \mathbb{R}^5) \cap C_b^{1,1,2}([0,t) \times \mathbb{R} \times \mathbb{R}^4)$ exists, then it is given by

$$f_t(s,v,u) := \mathbb{E}(\psi(X_t^{s,v,u})) \tag{D.8}$$

where $(X_r^{s,v,u} := (V_r, m_r, n_r, h_r, y_r) : r \in [s,t])$ is the unique (pathwise and in law) solution in $[s,t]$ of the stochastic differential equation:

$$\begin{aligned} V_r &= v + \int_s^r F(V_\theta, m_\theta, n_\theta, h_\theta) d\theta - \int_s^r J_{\mathrm{E}}(V_\theta - \alpha_\theta) ds - \int_s^t J_{\mathrm{Ch}} \beta_\theta (V_\theta - V_{\mathrm{rev}}) d\theta, \\ x_r &= u_x + \int_s^r \rho'_x(V_\theta)(1 - x_\theta) - \zeta'_x(V_\theta) x_\theta d\theta + \int_s^r \sigma'_x(V_\theta, x_\theta) dW_\theta^x, \quad x = m,n,h,y. \end{aligned}$$

Moreover, for v chosen in some fixed compact set, this solution is bounded independently of $s \in [0,t]$, and one has $x_r \in [0,1]$ for all $r \in [s,t]$. Hence, under the assumption that $\sigma > 0$, ρ_x and ζ_x are of class $C^2(\mathbb{R})$, one can moreover prove, following the lines of Friedman [22, p.124], that the function f_t defined by (D.8) actually is of class $C_b^{1,1,2}([0,t) \times \mathbb{R} \times \mathbb{R}^4)$ and solves the Cauchy problem (D.7). Putting $\psi = f_t$ in (D.6) yields

$$\int \psi(v,u) \nu_t(dv,du) = \int \mathbb{E}(\psi(X_t^{s,v,u})) \nu_0(dv,du)$$

for all $\psi \in C_0^2(\mathbb{R}^5)$, which uniquely determines ν_t. Notice that when $\sigma = 0$, the required regularity for ϕ and for f turns from $C^{1,1,2}$ to $C^{1,1,1}$ and the Feymann Kac formula in the argument can be replaced by the characteristics formula. The proof of part c) is complete.

d) This is immediate from parts b) and d) of Theorem D.1 . □

Proof of Corollary 2.7. Recall first that, for any $\nu \in \mathcal{P}_2(\mathbb{R} \times [0,1]^4)$ and $w \in \mathbb{R} \times [0,1]^4)$, one has

$$\mathcal{W}_2^2(\nu, \delta_w) = \int |z - w|^2 \nu(dz).$$

Moreover, for every $t \geq t_1$ and $N \geq 1$ it holds by exchangeability that:

$$\mathbb{E}\left(|X_t^{(i)} - \widehat{X}_t^{t_1,N}|^2\right) = \mathbb{E}\left(\mathcal{W}_2^2(\mu_t^N, \delta_{\widehat{X}_t^{t_1,N}})\right).$$

Therefore, it is enough to prove that, for any $t_1 \geq 0$,

$$\sup_{t_1 \leq t \leq t_1 + \delta} \mathbb{E}\left|\mathcal{W}_2^2(\mu_t, \delta_{\widehat{X}_t^{t_1,\infty}}) - \mathcal{W}_2^2(\mu_t^N, \delta_{\widehat{X}_t^{t_1,N}})\right| \to 0$$

as $N \to \infty$. Given $t \geq t_1$ and $N \geq 1$, let $\pi_t^N(dz,dz')$ be a coupling between μ_t and μ_t^N. Then, for some constant $C > 0$ not depending on $t \geq t_1$ nor on $N \geq 1$, we have

$$\begin{aligned} \left|\mathcal{W}_2^2(\mu_t, \delta_{\widehat{X}_t^{t_1,\infty}}) - \mathcal{W}_2^2(\mu_t^N, \delta_{\widehat{X}_t^{t_1,N}})\right| &= \left|\int \pi_t^N(dz,dz') \left[|z - \widehat{X}_t^{t_1,\infty}|^2 - |z' - \widehat{X}_t^{t_1,N}|^2\right]\right| \\ &\leq C\left[\int |z - z'| \pi_t^N(dz,dz') + |\widehat{X}_t^{t_1,\infty} - \widehat{X}_t^{t_1,N}|\right] \end{aligned}$$

since the supports of μ_t and μ_t^N and the processes $\widehat{X}_t^{t_1,\infty}$ and $\widehat{X}_t^{t_1,N}$ are uniformly bounded in $t \geq t_1$ and N. The latter property also allows us to write the dynamics in (2.6) and (2.12) using globally Lipschitz coefficients. Thanks to Gronwall's lemma this yields the estimates

$$\sup_{t_1 \leq t \leq t_1 + \delta} |\widehat{X}_t^{t_1,\infty} - \widehat{X}_t^{t_1,N}| \leq C_\delta |\widehat{X}_{t_1}^{t_1,\infty} - \widehat{X}_{t_1}^{t_1,N}| = C_\delta |\langle \mu_{t_1} \rangle - \langle \mu_{t_1}^N \rangle| \leq C_\delta \int |z - z'| \pi_{t_1}^N(dz,dz')$$

for some constant $C_\delta > 0$ not depending on N. Since $\int |z-z'|\pi_t^N(dz,dz') \le \left(\int |z-z'|^2\pi_t^N(dz,dz')\right)^{1/2}$, by taking the above couplings to be optimal for $\mathcal{W}_2$, we get the estimate

$$\sup_{t_1\le t\le t_1+\delta} \mathbb{E}\left|\mathcal{W}_2^2(\mu_t,\delta_{\widehat{X}_t^{t_1,\infty}}) - \mathcal{W}_2^2(\mu_t^N,\delta_{\widehat{X}_t^{t_1,N}})\right| \le C' \sup_{t_1\le t\le t_1+\delta} \mathbb{E}^{1/2}\left(\mathcal{W}_2^2(\mu_t,\mu_t^N)\right)$$

for some $C' > 0$. We conclude thanks to Theorem 2.5. □

E Strong Convergence Rate Result for the Exponential Projective Euler Scheme (EPES)

The main object of this section is to prove the convergence of the numerical scheme presented in Section 3 to the model (2.1) and establish the following rate of convergence

Proposition E.1. *Asumme Hypothesis 2.2, if $\chi(x) = O(x(1-x))$, then there exists a constant C depending on the parameters of the system, but independent of Δt, such that for any $i = 1,\dots,N$:*

$$\mathbb{E}\left[\left(V_t^{(i)} - \widehat{V}_t^{(i)}\right)^2\right] + \sum_{x=m,n,h,y} \mathbb{E}\left[|x_t^{(i)} - \widehat{x}_t^{(i)}|^2\right] \le C\Delta t.$$

We decompose the proof of this proposition in several preliminary results.

The next result follows from the uniform bound for $\widehat{V}_t^{(i)}$ (see iii) in Remark A.4) and some standard arguments on local approximation of SDEs, so we omit the proof.

Lemma E.2. *Under Hypothesis 2.2, there exists a constant C depending on the parameters of the system, but independent of Δt such that*

$$\sup_{i=1,\dots,N} \mathbb{E}\left[\left(\widehat{V}_t^{(i)} - \widehat{V}_{\eta(t)}^{(i)}\right)^2\right] \le C\Delta t^2, \qquad \sup_{i=1,\dots,N} \mathbb{E}\left[\left(\check{x}_t^{(i)} - \widehat{x}_{\eta(t)}^{(i)}\right)^2\right] \le C\Delta t.$$

Next we establish a the key step in the convergence of the scheme, namely that, with extremely high probability, the processes $\widehat{x}^{(i)}$ and $\check{x}^{(i)}$ coincide.

Lemma E.3. *Asumme Hypothesis 2.2, if $\chi(x) = O(x(1-x))$, then there exists a constant C depending on the parameters of the system, but independent of Δt, such that*

$$\sup_{i=1,\dots,N} \sum_{x=m,n,h,y} \mathbb{P}\left(\check{x}_t^{(i)} \notin [0,1]\right) \le \exp\left(-\frac{C}{\Delta t}\right).$$

Remark E.4. *It is not difficult to see that*

$$\begin{aligned}\mathbb{E}\left[\left(\check{x}_t^{(i)} - \widehat{x}_t^{(i)}\right)^2\right] &= \mathbb{E}\left[\left(\check{x}_t^{(i)} - \widehat{x}_t^{(i)}\right)^2 \mathbb{1}_{\{\check{x}_t^{(i)}\notin[0,1]\}}\right] \\ &\le \sqrt{\sup_{j=1,\dots,N} \mathbb{E}\left[2(\check{x}_t^j)^2+1\right] \mathbb{P}\left(\check{x}_t^{(i)} \notin [0,1]\right)} \le K\exp\left(-\frac{C}{2\Delta t}\right).\end{aligned}$$

Notice that the RHS above tends to zero faster than any power of Δt when $\Delta t \to 0$.

Proof of Lemma E.3. We first notice that conditional to $\mathcal{F}_{\eta(t)}$, $\check{x}^{(i)}$ corresponds to an Ornstein-Uhlenbeck process, therefore its law is Gaussian with known conditional mean and conditional variance given by

$$\begin{aligned}\mathbb{E}_{\eta(t)}\left[\check{x}_t^{(i)}\right] &= \widehat{x}_{\eta(t)}^{(i)} \exp\left(-\left(\rho_x+\zeta_x\right)(\widehat{V}_{\eta(t)}^{(i)})(t-\eta(t))\right) \\ &\quad + \frac{\rho_x(\widehat{V}_{\eta(t)}^{(i)})}{\left(\rho_x+\zeta_x\right)(\widehat{V}_{\eta(t)}^{(i)})}\left(1-\exp\left(-\left(\rho_x+\zeta_x\right)(\widehat{V}_{\eta(t)}^{(i)})(t-\eta(t))\right)\right), \\ \mathbb{V}\mathrm{ar}_{\eta(t)}\left[\check{x}_t^{(i)}\right] &= \frac{\sigma_x^2(\widehat{V}_{\eta(t)}^{(i)},\widehat{x}_{\eta(t)}^{(i)})}{2\left(\rho_x+\zeta_x\right)(\widehat{V}_{\eta(t)}^{(i)})}\left(1-\exp\left(-2\left(\rho_x+\zeta_x\right)(\widehat{V}_{\eta(t)}^{(i)})(t-\eta(t))\right)\right).\end{aligned}$$

Observe that the conditional variance is strictly positive if $t > \eta(t)$, $\widehat{x}^{(i)}_{\eta(t)} \neq 0$ and $\widehat{x}^{(i)}_{\eta(t)} \neq 1$. Since for $\widehat{x}^{(i)}_{\eta(t)} = 0$ or $\widehat{x}^{(i)}_{\eta(t)} = 1$ the diffusions coefficient vanish, and in that case the solution to the ODE for $\check{x}^{(i)}$ remains in $[0,1]$ almost surely, we can restrict ourselves to the case $\widehat{x}^{(i)}_{\eta(t)} \in (0,1)$.

Using the inequality for Gaussian concentration, conditional to $\mathcal{F}_{\eta(t)}$, we have

$$\mathbb{P}_{\eta(t)}\left(\check{x}^{(i)}_t \leq 0\right) = \mathbb{P}_{\eta(t)}\left(\frac{\check{x}^{(i)}_t - \mathbb{E}_{\eta(t)}\left[\check{x}^{(i)}_t\right]}{\sqrt{\mathbb{V}\mathrm{ar}_{\eta(t)}\left[\check{x}^{(i)}_t\right]}} \leq \frac{-\mathbb{E}_{\eta(t)}\left[\check{x}^{(i)}_t\right]}{\sqrt{\mathbb{V}\mathrm{ar}_{\eta(t)}\left[\check{x}^{(i)}_t\right]}}\right) \leq \frac{1}{2}\exp\left(-\frac{\mathbb{E}_{\eta(t)}\left[\check{x}^{(i)}_t\right]^2}{\mathbb{V}\mathrm{ar}_{\eta(t)}\left[\check{x}^{(i)}_t\right]}\right).$$

Since for t small enough

$$1 - \exp\left(-2\left(\rho_x + \zeta_x\right)(\widehat{V}^{(i)}_{\eta(t)})(t-\eta(t))\right) \leq 2\left(\rho_x + \zeta_x\right)(\widehat{V}^{(i)}_{\eta(t)})(t-\eta(t)),$$

and $t - \eta(t) \leq \Delta t$, we can bound the conditional variance, and then it follows

$$\mathbb{P}_{\eta(t)}\left(\check{x}^{(i)}_t \leq 0\right) \leq \frac{1}{2}\exp\left(-\frac{\mathbb{E}_{\eta(t)}\left[\check{x}^{(i)}_t\right]^2}{\sigma_x^2(\widehat{V}^{(i)}_{\eta(t)}, \widehat{x}^{(i)}_{\eta(t)})\Delta t}\right).$$

On the other hand, $\mathbb{E}_{\eta(t)}\left[\check{x}^{(i)}_t\right]$ is a weighted mean between to quantities in $[0,1]$, therefore

$$\mathbb{E}_{\eta(t)}\left[\check{x}^{(i)}_t\right] \geq \widehat{x}^{(i)}_{\eta(t)} \wedge \frac{\rho_x(\widehat{V}^{(i)}_{\eta(t)})}{\left(\rho_x + \zeta_x\right)(\widehat{V}^{(i)}_{\eta(t)})},$$

hence

$$\begin{aligned}\mathbb{P}_{\eta(t)}\left(\check{x}^{(i)}_t \leq 0\right) \leq \frac{1}{2}\exp\left(\frac{-\rho_x^2(\widehat{V}^{(i)}_{\eta(t)})}{\left(\rho_x + \zeta_x\right)^2(\widehat{V}^{(i)}_{\eta(t)})\sigma_x^2(\widehat{V}^{(i)}_{\eta(t)}, \widehat{x}^{(i)}_{\eta(t)})\Delta t}\right)\mathbb{1}_{\left\{\widehat{x}^{(i)}_{\eta(t)} \geq \frac{\rho_x(\widehat{V}^{(i)}_{\eta(t)})}{(\rho_x+\zeta_x)(\widehat{V}^{(i)}_{\eta(t)})}\right\}} \\ + \frac{1}{2}\exp\left(-\frac{\left(\widehat{x}^{(i)}_{\eta(t)}\right)^2}{\sigma_x^2(\widehat{V}^{(i)}_{\eta(t)}, \widehat{x}^{(i)}_{\eta(t)})\Delta t}\right)\mathbb{1}_{\left\{\widehat{x}^{(i)}_{\eta(t)} \leq \frac{\rho_x(\widehat{V}^{(i)}_{\eta(t)})}{(\rho_x+\zeta_x)(\widehat{V}^{(i)}_{\eta(t)})}\right\}}.\end{aligned} \tag{E.1}$$

To bound the first exponential in the right-hand side of the last inequality, we notice that since the process $\widehat{V}^{(i)}$ is uniformly bounded and σ is bounded, we can easily exhibit a constant $C_1 > 0$ independent of i such that

$$\frac{\rho_x^2(\widehat{V}^{(i)}_{\eta(t)})}{\left(\rho_x + \zeta_x\right)^2(\widehat{V}^{(i)}_{\eta(t)})\sigma_x^2(\widehat{V}^{(i)}_{\eta(t)}, \widehat{x}^{(i)}_{\eta(t)})} \geq C_1.$$

For the second term in the right-hand side of (E.1), since $x^2/\chi(x)^2$ is bounded from below in $(0,1)$, there exists $C_2 > 0$, such that

$$\frac{\left(\widehat{x}^{(i)}_{\eta(t)}\right)^2}{\sigma_x^2(\widehat{V}^{(i)}_{\eta(t)}, \widehat{x}^{(i)}_{\eta(t)})} \geq C_2,$$

from which we conclude

$$\mathbb{P}\left(\check{x}^{(i)}_t \leq 0\right) \leq \exp\left(-\frac{C_1 \wedge C_2}{\Delta t}\right).$$

An analogous computation shows

$$\mathbb{P}\left(\check{x}^{(i)}_t \geq 1\right) \leq \exp\left(-\frac{C_1 \wedge C_2}{\Delta t}\right).$$

□

The last preliminary step in the proof of Proposition E.1 is the following

Lemma E.5. *Under hypotheses of Proposition E.1, consider*

$$u(t) := \mathbb{E}\left[\left(V_t^{(i)} - \widehat{V}_t^{(i)}\right)^2\right] + \sum_{x=m,n,h,y} \mathbb{E}\left[|x_t^{(i)} - \check{x}_t^{(i)}|^2\right].$$

Then there exists a constant C depending on the parameters of the system, but independent of Δt, such that

$$u(t) \leq \left(u(\eta(t)) + C\Delta t^2\right) e^{C\Delta t}. \tag{E.2}$$

Proof. Thanks to the boundedness of the processes, drift and diffusion coefficients, that is b_x and σ_x, behave like Lipschitz functions, just as in the proof of Lemma A.1. Then, thanks to Itô formula and pivoting with in drift and diffusion with the point $(\widehat{V}_s^{(i)}, \check{x}_s^{(i)})$

$$\begin{aligned}
&\mathbb{E}\left[(x_t^{(i)} - \check{x}_t^{(i)})^2\right] \\
&\leq \mathbb{E}\left[(x_{\eta(t)}^{(i)} - \widehat{x}_{\eta(t)}^{(i)})^2\right] + 2\int_{\eta(t)}^t \mathbb{E}\left[(x_s^{(i)} - \check{x}_s^{(i)})\left(b_x(V_s^{(i)}, x_s^{(i)}) - b_x(\widehat{V}_s^{(i)}, \check{x}_s^{(i)})\right)\right] ds \\
&\quad + 2\int_{\eta(t)}^t \mathbb{E}\left[(x_s^{(i)} - \check{x}_s^{(i)})\left(b_x(\widehat{V}_s^{(i)}, \check{x}_s^{(i)}) - b_x(\widehat{V}_{\eta(t)}^{(i)}, \check{x}_s^{(i)})\right)\right] ds \\
&\quad + \int_{\eta(t)}^t 2\mathbb{E}\left[\left(\sigma_x(V_s^{(i)}, x_s^{(i)}) - \sigma_x(\widehat{V}_s^{(i)}, \check{x}_s^{(i)})\right)^2\right] + 2\mathbb{E}\left[\left(\sigma_x(\widehat{V}_s^{(i)}, \check{x}_s^{(i)}) - \sigma_x(\widehat{V}_{\eta(t)}^{(i)}, \widehat{x}_{\eta(t)}^{(i)})\right)^2\right] ds.
\end{aligned}$$

from where the Lipchitz property of the coefficients, Lemma E.2 to bound the terms involving the local error and some classical arguments lead to

$$\mathbb{E}\left[(x_t^{(i)} - \check{x}_t^{(i)})^2\right] \leq \mathbb{E}\left[(x_{\eta(t)}^{(i)} - \check{x}_{\eta(t)}^{(i)})^2\right] + C\int_{\eta(t)}^t \mathbb{E}\left[(x_s^{(i)} - \check{x}_s^{(i)})^2\right] + \mathbb{E}\left[(V_s^{(i)} - \widehat{V}_s^{(i)})^2\right] ds + C\Delta t^2.$$

On the other hand, for the voltage error we obtain first the a.s. bound

$$\begin{aligned}
\left(V_t^{(i)} - \widehat{V}_t^{(i)}\right)^2 \leq \left(V_{\eta(t)}^{(i)} - \widehat{V}_{\eta(t)}^{(i)}\right)^2 &+ C\int_{\eta(t)}^t |V_s^{(i)} - \widehat{V}_s^{(i)}|^2 + \sum_{x=m,n,h} |x_s^{(i)} - \widehat{x}_{\eta(t)}^{(i)}|^2 ds \\
&+ C\int_{\eta(t)}^t |V_s^{(i)} - \widehat{V}_s^{(i)}|^2 + \frac{1}{N}\sum_{j=1}^N |V_s^{(j)} - \widehat{V}_{\eta(t)}^{(j)}|^2 ds, \\
&+ C\int_{\eta(t)}^t |V_s^{(i)} - \widehat{V}_s^{(i)}|^2 + (V_s^{(i)} + V_{\text{rev}})^2 \frac{1}{N}\sum_{j=1}^N |y_s^{(j)} - \widehat{y}_{\eta(t)}^{(j)}|^2 ds.
\end{aligned}$$

Thanks to the exchangeability of the particles, it follows that

$$\mathbb{E}\left[\frac{1}{N}\sum_{j=1}^N |y_s^{(j)} - \widehat{y}_{\eta(t)}^{(j)}|^2\right] = \mathbb{E}\left[\left(y_s^{(i)} - \widehat{y}_{\eta(t)}^{(i)}\right)^2\right], \quad \mathbb{E}\left[\frac{1}{N}\sum_{j=1}^N |V_s^{(j)} - \widehat{V}_{\eta(t)}^{(j)}|^2\right] = \mathbb{E}\left[\left(V_s^{(i)} - \widehat{V}_{\eta(t)}^{(i)}\right)^2\right],$$

and then, since the processes are uniformly bounded, we get that

$$
\begin{aligned}
&\mathbb{E}\left[\left(V_t^{(i)}-\widehat{V}_t^{(i)}\right)^2\right]\\
&\quad\leq \mathbb{E}\left[\left(V_{\eta(t)}^{(i)}-\widehat{V}_{\eta(t)}^{(i)}\right)^2\right]+C\int_{\eta(t)}^t \mathbb{E}\left[\left(V_s^{(i)}-\widehat{V}_s^{(i)}\right)^2\right]ds+\\
&\qquad+C\int_{\eta(t)}^t \mathbb{E}\left[\left(V_s^{(i)}-\widehat{V}_{\eta(t)}^{(i)}\right)^2\right]+\sum_{x=m,n,h,y}\mathbb{E}\left[|x_s^{(i)}-\widehat{x}_{\eta(t)}^{(i)}|^2\right]ds\\
&\quad\leq \mathbb{E}\left[\left(V_{\eta(t)}^{(i)}-\widehat{V}_{\eta(t)}^{(i)}\right)^2\right]+C\int_{\eta(t)}^t \mathbb{E}\left[|V_s^{(i)}-\widehat{V}_s^{(i)}|^2\right]ds+\\
&\qquad+C\int_{\eta(t)}^t \mathbb{E}\left[\left(V_s^{(i)}-\widehat{V}_s^{(i)}\right)^2\right]+C\Delta t^2+\sum_{x=m,n,h,y}\mathbb{E}\left[|x_s^{(i)}-\check{x}_s^{(i)}|^2\right]+C\Delta t ds\\
&\quad\leq \mathbb{E}\left[\left(V_{\eta(t)}^{(i)}-\widehat{V}_{\eta(t)}^{(i)}\right)^2\right]+C\int_{\eta(t)}^t \mathbb{E}\left[\left(V_s^{(i)}-\widehat{V}_s^{(i)}\right)^2\right]+\sum_{x=m,n,h,y}\mathbb{E}\left[|x_s^{(i)}-\check{x}_s^{(i)}|^2\right]ds+C\Delta t^2.
\end{aligned}
$$

We can summarize the previous computations as

$$
u(t)\leq u(\eta(t))+C\int_{\eta(t)}^t u(s)ds+C\Delta t^2,
$$

from where we conclude thanks to Gronwall's inequality. □

Proof of Proposition E.1 . From the previous Lemma, denoting $u_k=u(t_k)$ we obtain the following recurrence relationship:

$$
u_{k+1}\leq Au_k+B,\quad u_0=0,\quad A=e^{C\Delta t},\quad B=C\Delta t^2e^{C\Delta t}.
$$

Iterating this inequality, it is easy to conclude that

$$
u_{k+1}\leq B\frac{A^{k+1}-1}{A-1}\leq C\Delta t^2\frac{\left(e^{C\Delta t}\right)^{k+1}}{e^{C\Delta t}-1}=C\Delta t^2\frac{e^{Ct_{k+1}}}{e^{C\Delta t}-1}.
$$

But when $\Delta t\to 0$, we have $e^{C\Delta t}-1\sim C\Delta t$, and therefore $u_{k+1}\leq C\Delta t$. Inserting this in (E.2), we conclude

$$
\begin{aligned}
&\mathbb{E}\left[\left(V_t^{(i)}-\widehat{V}_t^{(i)}\right)^2\right]+\sum_{x=m,n,h,y}\mathbb{E}\left[|x_t^{(i)}-\widehat{x}_t^{(i)}|^2\right]\\
&\quad\leq \mathbb{E}\left[\left(V_t^{(i)}-\widehat{V}_t^{(i)}\right)^2\right]+\sum_x\mathbb{E}\left[|x_t^{(i)}-\check{x}_t^{(i)}|^2\right]+\mathbb{E}\left[|\check{x}_t^{(i)}-\widehat{x}_t^{(i)}|^2\right]\leq C\Delta t+\sum_x\mathbb{P}\left(\check{x}_t^{(i)}\notin[0,1]\right),
\end{aligned}
$$

from where the statement follows, applying Lemma E.3.

□

References

[1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008. 18

[2] T.D. Austin. The emergence of the deterministic hodgkin–huxley equations as a limit from the underlying stochastic ion-channel mechanism. *The Annals of Applied Probability*, 18(4):1279–1325, 2008. 2

[3] N. Axmacher, F. Mormann, G. Fernández, C. E. Elger, and J. Fell. Memory formation by neuronal synchronization. *Brain Research Reviews*, 52(1):170 – 182, 2006. 3

[4] J. Baladron, D. Fasoli, O. Faugeras, and J. Touboul. Mean field description of and propagation of chaos in networks of hodgkin-huxley and fitzhugh-nagumo neurons. *Journal of Mathematical Neuroscience*, 2(10), 2012. 3

[5] Nils Berglund and Barbara Gentz. On the noise-induced passage through an unstable periodic orbit i: Two-level model. *Journal of statistical physics*, 114(5-6):1577–1618, 2004. 3

[6] Nils Berglund and Barbara Gentz. On the noise-induced passage through an unstable periodic orbit ii: General case. *SIAM Journal on Mathematical Analysis*, 46(1):310–352, 2014. 3

[7] L. Bertini, G. Giacomin, and K. Pakdaman. Dynamical aspects of mean field plane rotators and the kuramoto model. *Journal of Statistical Physics*, 138(1):270–290, 2010. 3

[8] L. Bertini, G. Giacomin, and C. Poquet. Synchronization and random long time dynamics for mean-field plane rotators. *Probab. Theory Related Fields*, 160(3-4):593–653, 2014. 16

[9] M. Bossy, J. Espina, J. Morice, C. Paris, and A. Rosseau. Modeling the wind circulation around mills with a lagrangian stochastic approach. *SMAI Journal of Computational Mathematics*, 2:177–214, 2016. 9

[10] M. Bossy, O. Faugeras, and D. Talay. Clarification and complement to "mean-field description and propagation of chaos in networks of hodgkin–huxley and fitzhugh–nagumo neurons". *The Journal of Mathematical Neuroscience (JMN)*, 5(1):1–23, 2015. 2, 3, 5, 14, 17, 28

[11] P. C. Bressloff and Y. M. Lai. Stochastic synchronization of neuronal populations with intrinsic and extrinsic noise. *The Journal of Mathematical Neuroscience*, 1(1):2, May 2011. 3

[12] A. N. Burkitt. A review of the integrate-and-fire neuron model: I. homogeneous synaptic input. *Biological Cybernetics*, 95(1):1–19, Jul 2006. 2

[13] A. N. Burkitt. A review of the integrate-and-fire neuron model: Ii. inhomogeneous synaptic input and network properties. *Biological Cybernetics*, 95(2):97–112, Aug 2006. 2

[14] T. Chan, G. Golub, and R. LeVeque. Algorithms for computing the sample variance: Analysis and recommendations. *The American Statistician*, 37(3):242–247, 1983. 10

[15] C. E. Dangerfield, D. Kay, and K. Burrage. Modeling ion channel dynamics through reflected stochastic differential equations. *Phys. Rev. E*, 85:051907, May 2012. 2

[16] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré. Global solvability of a networked integrate-and-fire model of mckean–vlasov type. *The Annals of Applied Probability*, 25(4):2096–2133, 2015. 3

[17] G. B. Ermentrout and D. H. Terman. *Mathematical Foundations of Neuroscience*. Springer-Verlag New York, 2010. 2, 3, 4, 5, 9, 10, 14

[18] O. Faugeras, J. Touboul, and B. Cessac. A constructive mean-field analysis of multi population neural networks with random synaptic weights and stochastic inputs. *Frontiers in Computational Neuroscience*, 3:1, 2009. 3

[19] R. FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1(6):445–466, 1961. 2

[20] N. Fournier and A. Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3-4):707–738, 2015. 8, 29

[21] N. Fournier and E. Löcherbach. On a toy model of interacting neurons. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 52(4):1844–1876, 11 2016. 3

[22] A. Friedman. *Stochastic differential equations and applications*. Dover Publications Inc., Mineola, NY, 2006. Two volumes bound as one, Reprint of the 1975 and 1976 original published in two volumes. 30

[23] J. Gärtner. On the McKean-Vlasov limit for interacting diffusions. *Mathematische Nachrichten*, 137:197–248, 1988. 29

[24] G. Giacomin, E. Luçon, and C. Poquet. Coherence stability and effect of random natural frequencies in populations of coupled oscillators. *J. Dynam. Differential Equations*, 26(2):333–367, 2014. 16

[25] J. Goldwyn, Nikita S. Imennov, M. Famulare, and E. Shea-Brown. Stochastic differential equation models for ion channel noise in hodgkin-huxley neurons. *Physical Review E*, 83(4):041908, 2011. 2

[26] J. Goldwyn and E. Shea-Brown. The what and where of adding channel noise to the hodgkin-huxley equations. *PLoS Computational Biology*, 7(11):e1002247, 2011. 2

[27] D. Hansel and G. Mato. Patterns of synchrony in a heterogeneous hodgkin-huxley neural network with weak coupling. *Physica A: Statistical Mechanics and its Applications*, 200(1-4):662–669, 1993. 3, 16

[28] D. Hansel, G. Mato, and C. Meunier. Phase dynamics for weakly coupled hodgkin-huxley neurons. *EPL (Europhysics Letters)*, 23(5):367, 1993. 3, 16

[29] A. Hodgkin and A. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*, 117, 1952. 1, 2, 5

[30] S. G. Hormuzdi, M. A. Filippov, G. Mitropoulou, H. Monyer, and R. Bruzzone. Electrical synapses: a dynamic signaling system that shapes the activity of neuronal networks. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1662(1–2):113 – 137, 2004. The Connexins. 4

[31] E. M. Izhikevich. *Dynamical Systems in Neuroscience*. The MIT Press, 2007. 2, 3

[32] P. Jiruska, M. de Curtis, J. G. R. Jefferys, C. A. Schevon, S. J. Schiff, and K. Schindler. Synchronization and desynchronization in epilepsy: controversies and hypotheses. *The Journal of Physiology*, 591(4):787–797, 2013. 3

[33] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer, 2nd edition, 1991. 30

[34] Nancy Kopell and Bard Ermentrout. Chemical and electrical synapses perform complementary roles in the synchronization of interneuronal networks. *Proceedings of the National Academy of Sciences*, 101(43):15482–15487, 2004. 14

[35] Y. Kuramoto. *Chemical Oscillations, Waves, and Turbulence*. Springer Berlin Heidelberg, 1984. 3

[36] L. Lapicque. Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarization. *J Physiol Pathol Gen (Paris)*, 9:620–635, 1907. 2

[37] E. Luçon and C. Poquet. Long time dynamics and disorder-induced traveling waves in the stochastic Kuramoto model. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(3):1196–1240, 2017. 16

[38] Sashi Marella and G Bard Ermentrout. Class-ii neurons display a higher degree of stochastic synchronization than class-i neurons. *Physical review E*, 77(4):041918, 2008. 3

[39] S. Méléard. Asymptotic behaviour of some interacting particle systems; mckean-vlasov and boltzmann models. In *Probabilistic models for nonlinear partial differential equations*, pages 42–95. Springer, 1996. 3, 8, 26, 27, 28

[40] S. Mischler, C. Quiñinao, and J. Touboul. On a kinetic fitzhugh–nagumo model of neuronal network. *Communications in Mathematical Physics*, 342(3):1001–1042, 2016. 3

[41] C. Morris and H. Lecar. Voltage oscillations in the barnacle gian muscle fiber. *Biophysical Journal.*, 31(1):193–213, 1981. 2

[42] J. Nagumo, S. Arimoto, and S. Yoshizawa. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE*, 1962. 2

[43] S. Ostojic, N. Brunel, and V. Hakim. Synchronization properties of networks of electrically coupled neurons in the presence of noise and heterogeneities. *Journal of Computational Neuroscience*, 26(3):369, Nov 2008. 3

[44] K. Pakdaman, M. Thieullen, and G. Wainrib. Fluid limit theorems for stochastic hybrid systems with aplication to neuron models. *Adv. in Appl. Probab.*, 42(3):761–794, 09 2010. 2, 5

[45] B. Perthame and D. Salort. On a voltage-conductance kinetic system for integrate and fire neural networks. *Kinetic and Related Models*, 6(4):841–864, 2013. 3

[46] AS Pikovskii. Synchronization and stochastization of nonlinear oscillations by external noise. In *Nonlinear and Turbulent Processes in Physics*, volume 1, page 1601, 1984. 3

[47] Arkady Pikovsky, Michael Rosenblum, and Jürgen Kurths. *Synchronization: a universal concept in nonlinear sciences*, volume 12. Cambridge university press, 2003. 3

[48] L. Sacerdote and M. Giraudo. *Stochastic Integrate and Fire Models: A Review on Mathematical Methods and Their Applications*, pages 99–148. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. 2

[49] A.-S. Sznitman. Topics in propagation of chaos. In *Ecole d'été de probabilités de Saint-Flour XIX—1989*, pages 165–251. Springer, 1991. 3, 8, 26, 28

[50] C. Villani. *Optimal transport, old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. 7

[51] G. Wainrib. *Randomness in neurons: a multiscale probabilistic analysis*. PhD thesis, École Polytechnique, 2010. 2, 3